**Explore natural language processing**

Natural language processing is the branch of artificial intelligence that works on enabling applications to  understand and use text and spoken words.

The goal of natural language processing is to ensure AI can interpret and operate with written and spoken language as efficiently as humans. The input and output devices of our computers and other technology are often barriers to maximizing benefit. Keyboards, pads, pointers, controllers, screens, and printers are mostly just substituting for the richer tools of speaking and hearing.

What are the advantages of using natural language processing in our current technologies?

Does the natural language processing of text and speech pose any dangers to our world?

List a few examples of how language processing can help you or your organization.

As you continue your research you may also like to investigate the use of natural language processing in AI for these, and other, use cases:

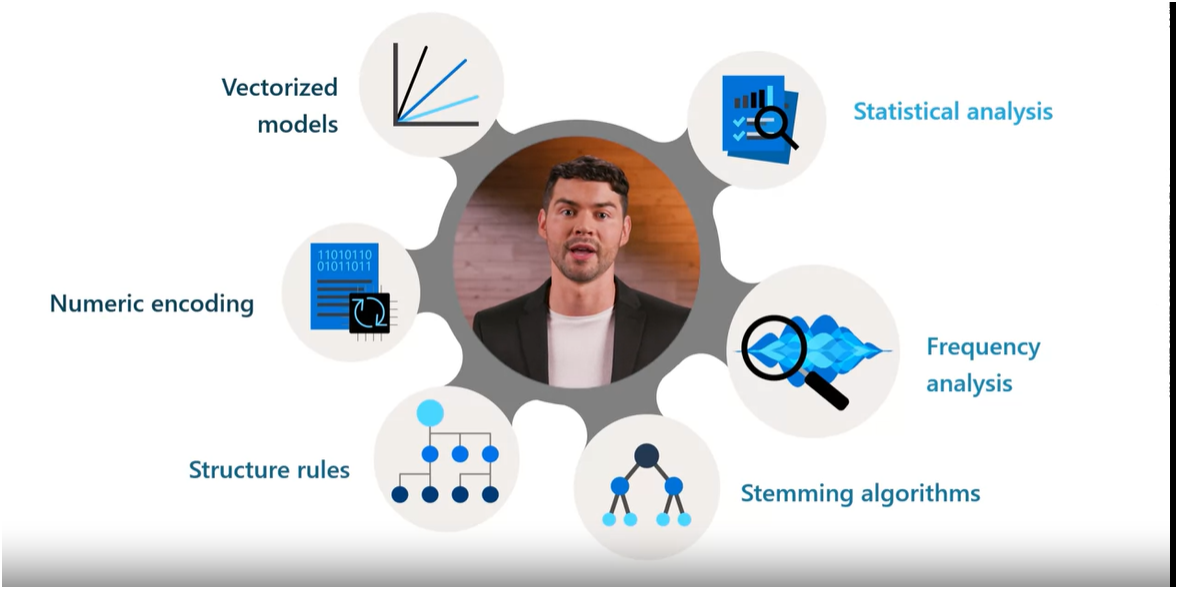
* Virtual assistants and chatbots
* Detecting anti-social activity on social media
* National security
* Spam detection
* Accessibility to support social inclusion
* Targeted marketing

Share your thoughts with your classmates and see what other people think about the current and future benefits or dangers of natural language processing in your personal and business lives.

# What is Text Analytics?



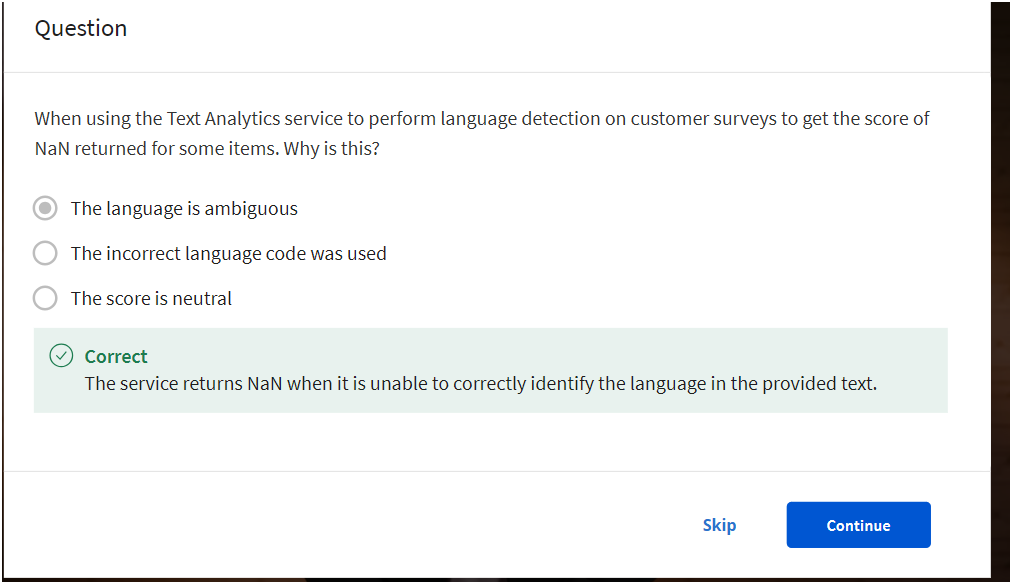
1. Text analysis: Text analysis is the process of evaluating different aspects of a document or phrase to gain insights into its content. Humans can read and understand text without considering grammar rules, and they can identify key phrases, names, and even the sentiment behind the text. Text analytics is the process of using AI algorithms to evaluate these attributes in text.
2. Techniques used in text analysis: There are several techniques used to build text analyzing software, including statistical analysis of terms, frequency analysis, stemming or lemmatization algorithms, linguistic structure rules, and encoding words as numeric features.
3. Text analytics cognitive service in Microsoft Azure: Microsoft Azure provides a text analytics cognitive service that simplifies application development by using pre-trained models. This service can determine the language of a document, perform sentiment analysis to determine positive or negative sentiment, and extract key phrases and entities from text.
4. Applications of text analytics: Text analytics can be applied to various applications, such as social media feed analyzers to detect sentiment around a political campaign or product, document search applications to extract key phrases for summarization, and tools to extract brand information or company names from text for identification purposes.



What are some techniques used in text analysis to build text analyzing software?

1. Statistical analysis of terms: This involves analyzing the frequency of terms used in the text. It includes removing common stop words (e.g., "a" or "the") and performing frequency analysis of the remaining words to gain insights into the main subject of the text.
2. Stemming or lemmatization algorithms: These techniques normalize words by reducing them to their base or root form. For example, words like "power," "powered," and "powerful" would be interpreted as the same word.
3. Linguistic structure rules: This technique involves analyzing the structure of sentences by breaking them down into tree-like structures. For example, a noun phrase might contain nouns, verbs, and adjectives, which helps in understanding the relationships between different parts of the text.
4. Encoding words as numeric features: This technique involves representing words and terms as numeric features that can be used to train machine learning models. For example, a text document can be classified based on the terms it contains, such as positive or negative sentiment.
5. Creating vectorized models: This technique captures semantic relationships between words by assigning them locations in n-dimensional space. For example, words like "flower" and "plant" might be assigned values that locate them close to each other, while "skateboard" might be positioned further away.

# Get started with Text Analytics on Azure



# Analyze text with Language Studio

Read the instruction from the reading file and setup Microsoft Azure Language Studio



Copy past the text given below :

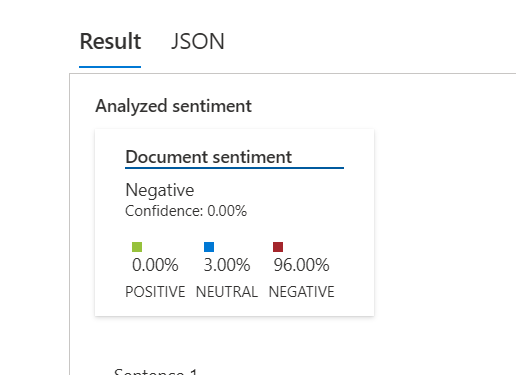
Tired hotel with poor service

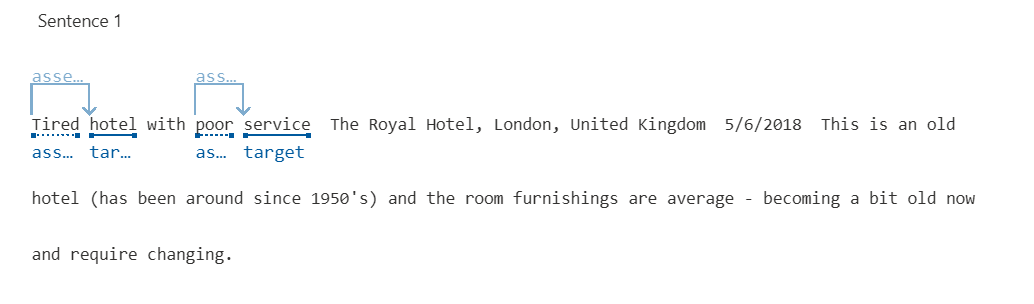
The Royal Hotel, London, United Kingdom

5/6/2018

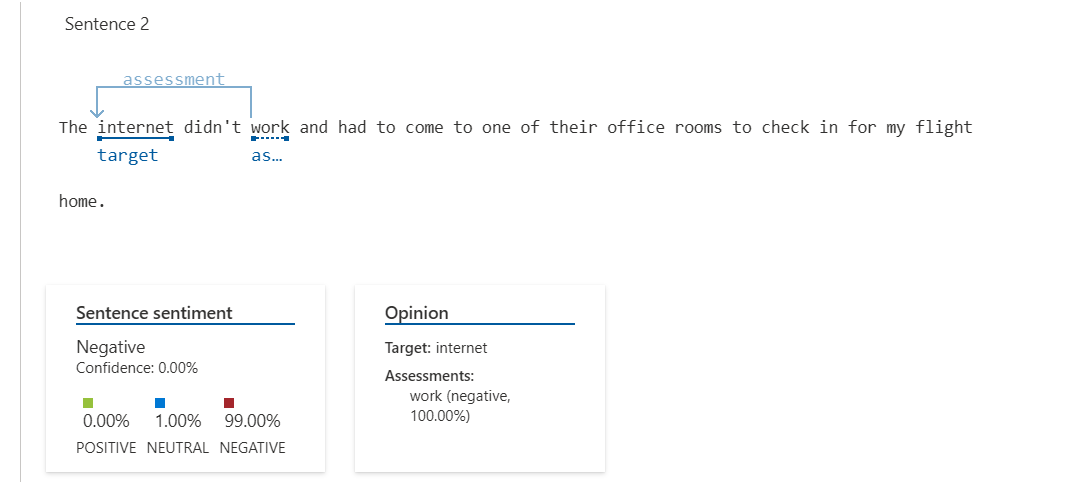
This is an old hotel (has been around since 1950's) and the room furnishings are average - becoming a bit old now and require changing. The internet didn't work and had to come to one of their office rooms to check in for my flight home. The website says it's close to the British Museum, but it's too far to walk.

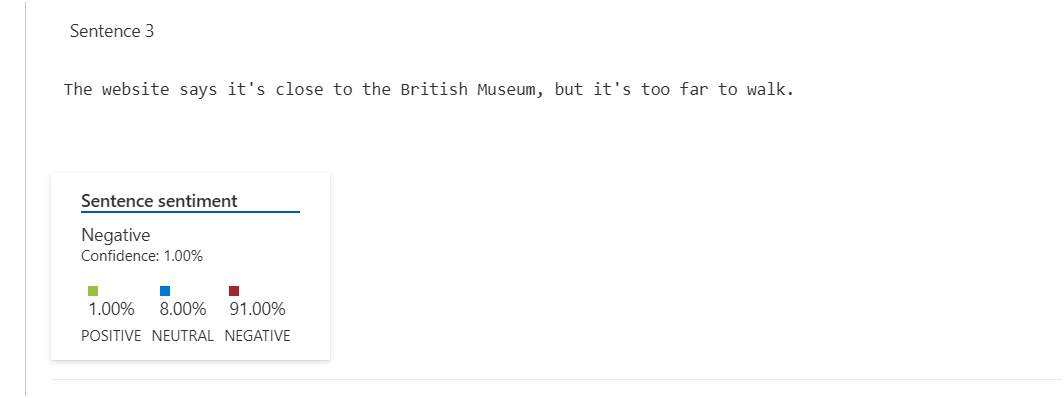
Result

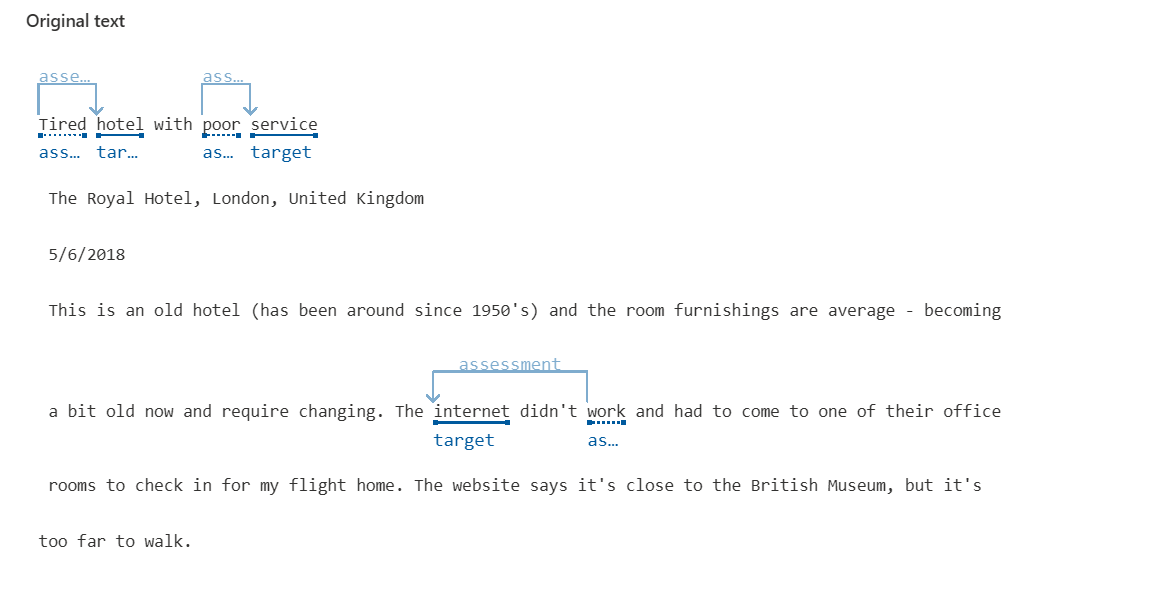












Copy past the text given below :

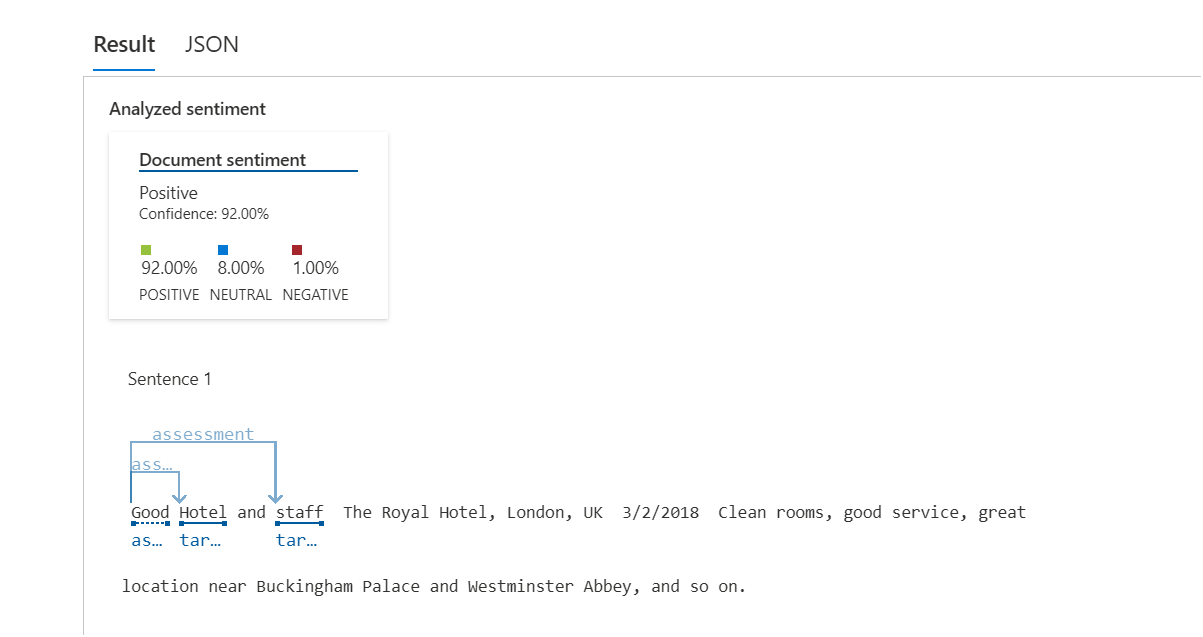
Good Hotel and staff

The Royal Hotel, London, UK

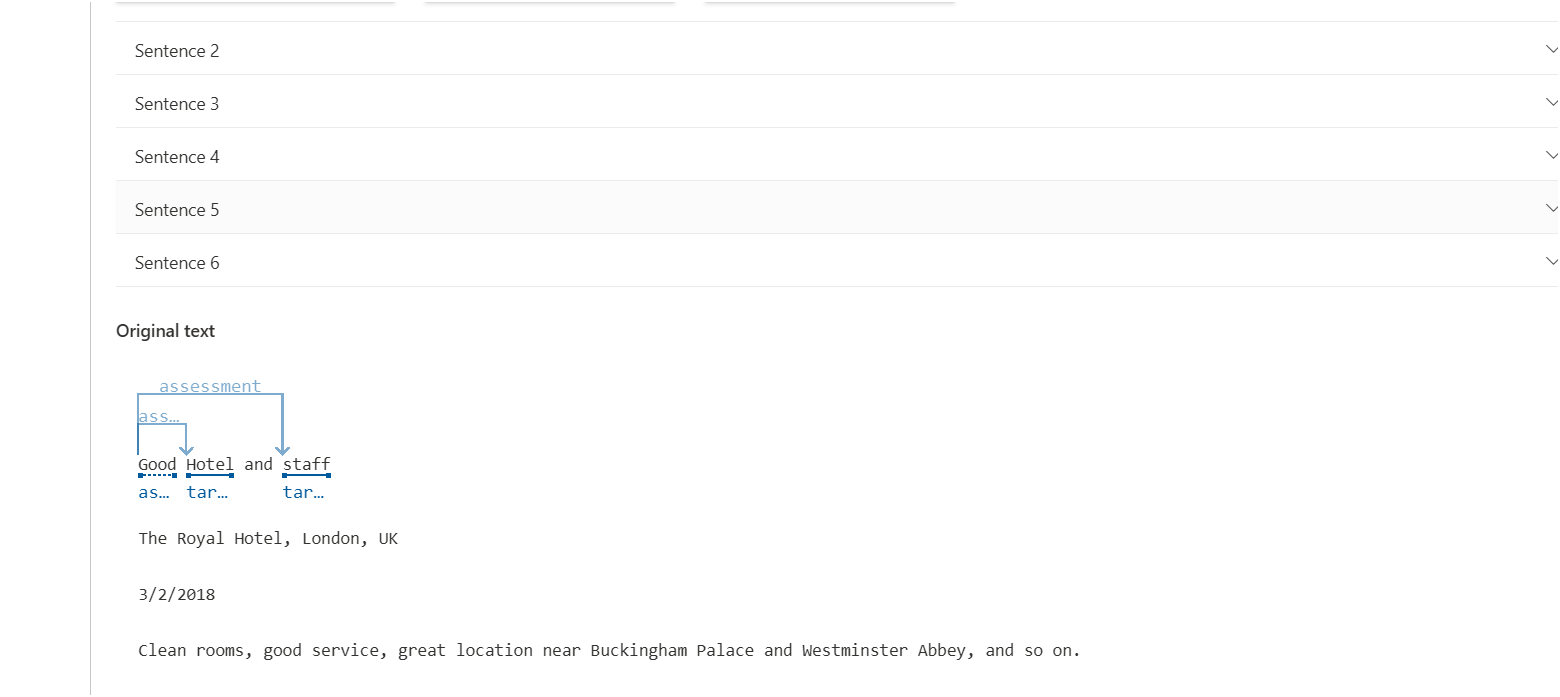
3/2/2018

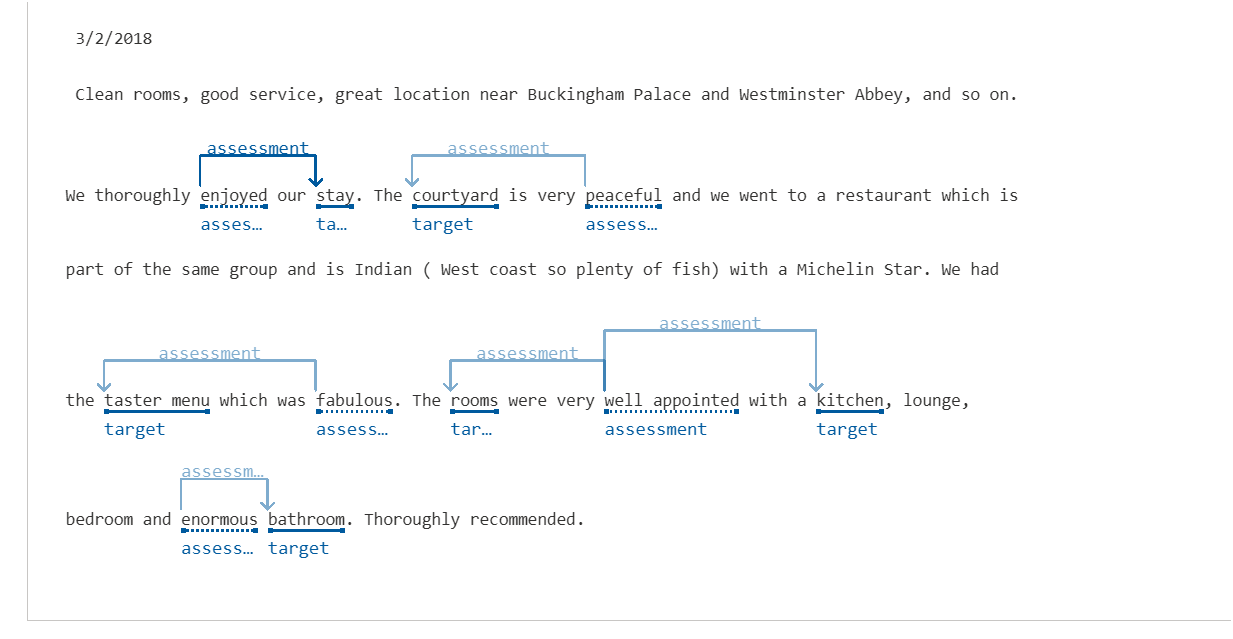
Clean rooms, good service, great location near Buckingham Palace and Westminster Abbey, and so on. We thoroughly enjoyed our stay. The courtyard is very peaceful and we went to a restaurant which is part of the same group and is Indian ( West coast so plenty of fish) with a Michelin Star. We had the taster menu which was fabulous. The rooms were very well appointed with a kitchen, lounge, bedroom and enormous bathroom. Thoroughly recommended.

Result:







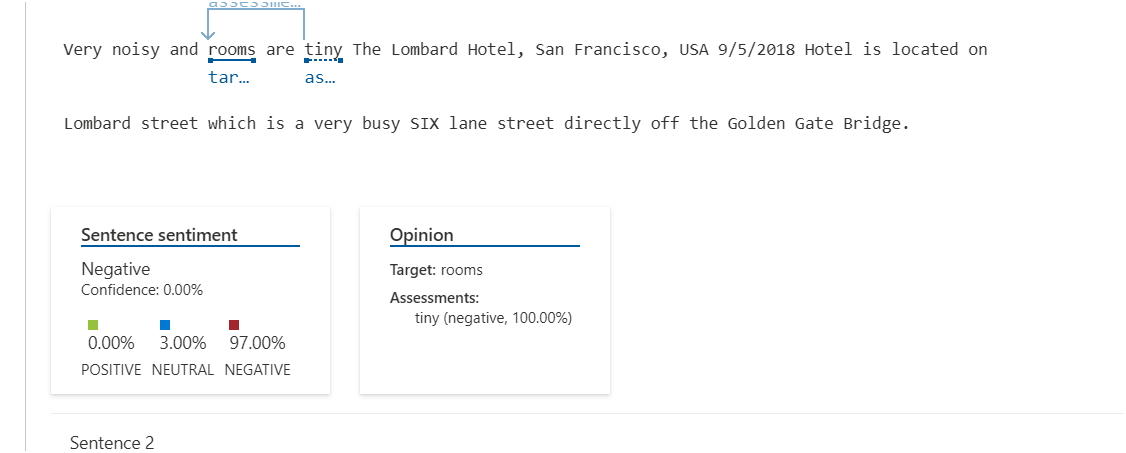


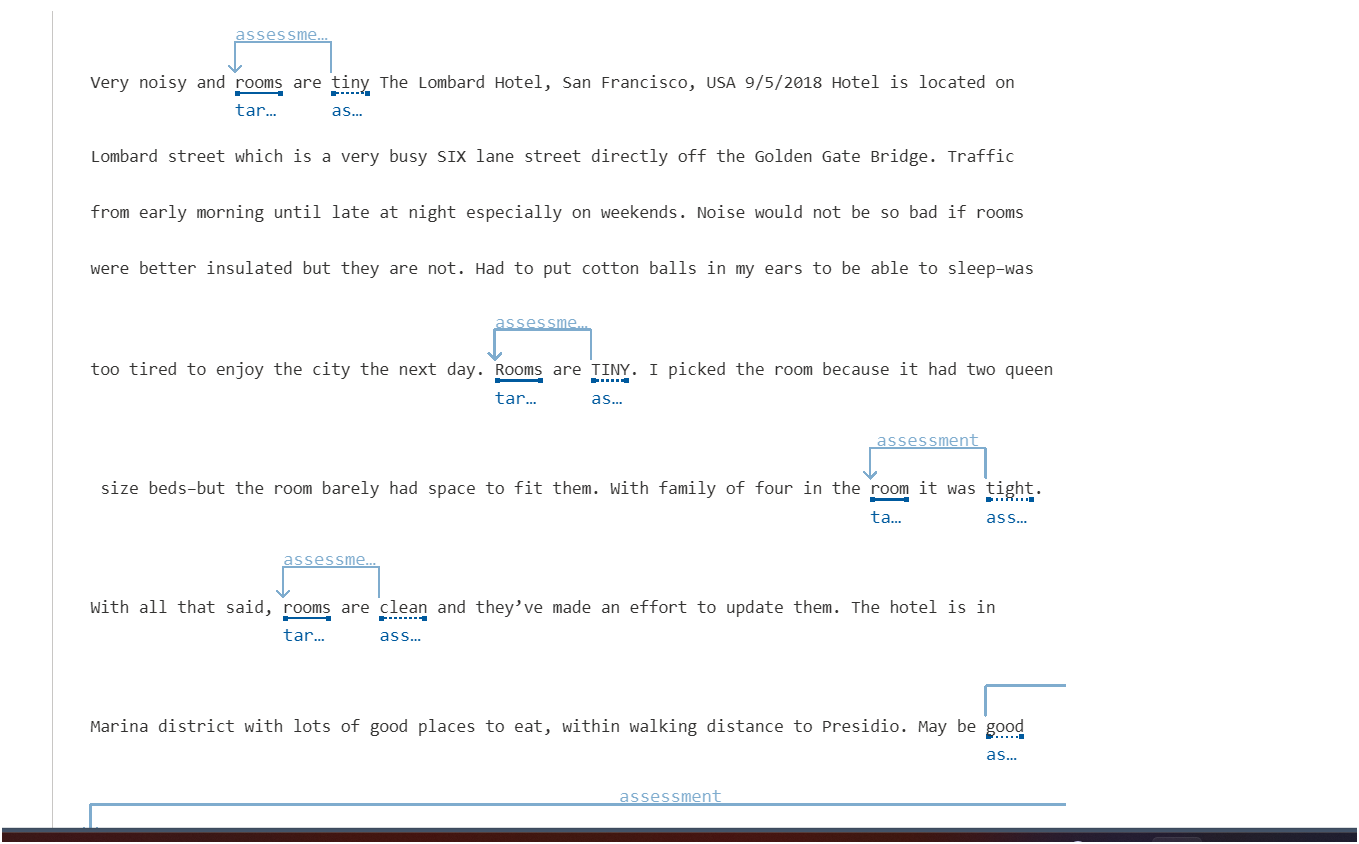
Copy past the text given below :

Very noisy and rooms are tiny The Lombard Hotel, San Francisco, USA 9/5/2018 Hotel is located on Lombard street which is a very busy SIX lane street directly off the Golden Gate Bridge. Traffic from early morning until late at night especially on weekends. Noise would not be so bad if rooms were better insulated but they are not. Had to put cotton balls in my ears to be able to sleep–was too tired to enjoy the city the next day. Rooms are TINY. I picked the room because it had two queen size beds–but the room barely had space to fit them. With family of four in the room it was tight. With all that said, rooms are clean and they’ve made an effort to update them. The hotel is in Marina district with lots of good places to eat, within walking distance to Presidio. May be good hotel for young stay-up-late adults on a budget

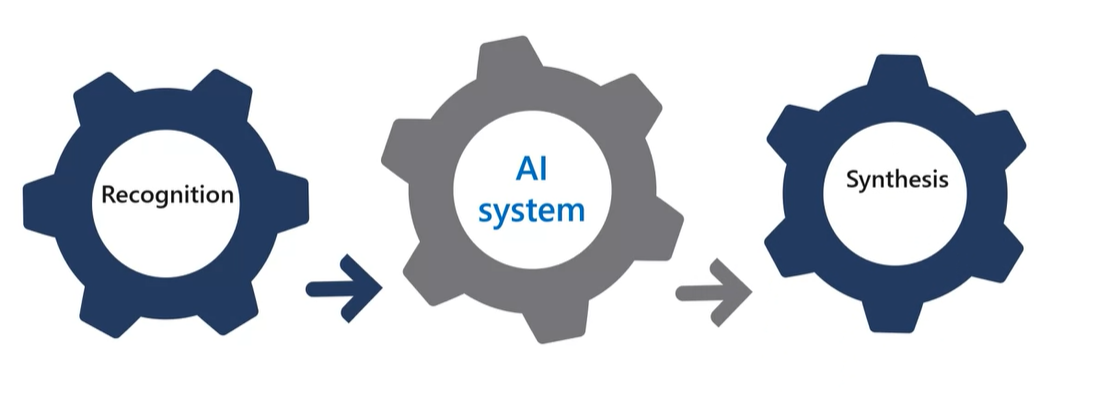
Results:







# Recognize and synthesize speech



* The video lecture introduces the importance of speech recognition and synthesis in enabling AI systems to accept vocal commands and provide spoken responses.
* Speech recognition involves converting spoken words into data that can be processed. This can be done by analyzing speech patterns and mapping them to words using acoustic and language models.
* The video explains that speech recognition can be used for various purposes, such as providing closed captions for videos, transcribing phone calls or meetings, automated note dictation, and determining user input for further processing.
* Speech synthesis, on the other hand, involves converting text into spoken speech. This process includes tokenizing the text, assigning phonetic sounds to words, and generating audio waveforms using a chosen voice.
* The video highlights the applications of speech synthesis, such as generating spoken responses to user input, creating voicemail messages, reading email or text messages aloud, and broadcasting announcements in public locations.
* Microsoft Azure Cognitive Services provides tools and services for speech recognition and synthesis, making it easier for developers to incorporate these capabilities into their applications.
* By leveraging Azure Cognitive Services, developers can build AI systems that can understand and respond to spoken language, enhancing user experiences and enabling more natural interactions.

How does speech recognition convert spoken words into data that can be processed?

Speech recognition converts spoken words into data that can be processed through a series of steps:

1. Audio Input: The speech recognition system takes in audio input, which can be in the form of a recorded voice in an audio file or live audio from a microphone.
2. Speech Analysis: The audio input is analyzed to identify speech patterns and distinguish them from background noise. This analysis involves techniques such as signal processing, filtering, and feature extraction.
3. Acoustic Modeling: An acoustic model is used to convert the audio signal into phonemes, which are representations of specific sounds in a language. The acoustic model maps the audio features extracted from the speech analysis to these phonemes.
4. Language Modeling: A language model is employed to map the phonemes to words. It uses statistical algorithms to predict the most probable sequence of words based on the phonemes detected by the acoustic model. The language model helps in improving the accuracy of recognizing words and understanding the context.
5. Text Representation: The recognized words are typically converted into a text representation, which can be further processed and analyzed. This text representation allows for various applications, such as closed captions for videos, transcription of phone calls or meetings, automated note dictation, and determining user input for further processing.

By following these steps, speech recognition systems can convert spoken words into data that can be processed and utilized by AI systems for a wide range of applications.

# Get Started with Speech on Azure

In this video lecture, the instructor discusses the speech recognition and speech synthesis capabilities offered by Microsoft Azure's Speech Cognitive Service. The video covers the following topics:

* Introduction to the Speech-to-Text API and the Text-to-Speech API provided by Microsoft Azure's Speech Cognitive Service.
* Provisioning the appropriate resource in your Azure subscription to use the speech service.
* Two types of resources available: Speech Resource and Cognitive Services Resource.
* Real-time and batch transcription of audio into text using the Speech-to-Text API.
* The Universal Language Model used by the Speech-to-Text API, which is trained by Microsoft and optimized for conversational and dictation scenarios.
* Creating and training custom models for speech-to-text.
* Real-time speech-to-text transcription for scenarios like presentations or demos.
* Batch transcription for audio recordings stored on file shares or Azure Storage.
* Asynchronous execution of batch transcription jobs.
* Converting text input to audible speech using the Text-to-Speech API.
* Specifying the voice to be used for speech synthesis, including predefined voices and neural voices.
* Developing custom voices for use with the Text-to-Speech API.
* Language support for both the Speech-to-Text and Text-to-Speech APIs.

# Exercise - Use the Speech service

**Introduction**

Completed100 XP

* 1 minute

AI speech capabilities enable us to manage home and auto systems with voice instructions, get answers from computers for spoken questions, generate captions from audio, and much more.

To enable this kind of interaction, the AI system must support two capabilities:

* **Speech recognition** - the ability to detect and interpret spoken input
* **Speech synthesis** - the ability to generate spoken output

**Azure AI Speech** provides speech to text and text to speech capabilities through speech recognition and synthesis. You can use prebuilt and custom Speech service models for a variety of tasks, from transcribing audio to text with high accuracy, to identifying speakers in conversations, creating custom voices, and more. Next you'll learn how AI speech capabilities work.

# Understand speech recognition and synthesis

**Speech recognition** takes the spoken word and converts it into data that can be processed - often by transcribing it into text. The spoken words can be in the form of a recorded voice in an audio file, or live audio from a microphone. Speech patterns are analyzed in the audio to determine recognizable patterns that are mapped to words. To accomplish this, the software typically uses multiple models, including:

* An *acoustic* model that converts the audio signal into phonemes (representations of specific sounds).
* A *language* model that maps phonemes to words, usually using a statistical algorithm that predicts the most probable sequence of words based on the phonemes.

The recognized words are typically converted to text, which you can use for various purposes, such as:

* Providing closed captions for recorded or live videos
* Creating a transcript of a phone call or meeting
* Automated note dictation
* Determining intended user input for further processing

**Speech synthesis** is concerned with vocalizing data, usually by converting text to speech. A speech synthesis solution typically requires the following information:

* The text to be spoken
* The voice to be used to vocalize the speech

To synthesize speech, the system typically *tokenizes* the text to break it down into individual words, and assigns phonetic sounds to each word. It then breaks the phonetic transcription into *prosodic* units (such as phrases, clauses, or sentences) to create phonemes that will be converted to audio format. These phonemes are then synthesized as audio and can be assigned a particular voice, speaking rate, pitch, and volume.

You can use the output of speech synthesis for many purposes, including:

* Generating spoken responses to user input
* Creating voice menus for telephone systems
* Reading email or text messages aloud in hands-free scenarios
* Broadcasting announcements in public locations, such as railway stations or airports

# Get started with speech on Azure

Microsoft Azure offers both speech recognition and speech synthesis capabilities through **Azure AI Speech** service, which includes the following application programming interfaces (APIs):

* The **Speech to text** API
* The **Text to speech** API

## Azure resources for Azure AI Speech

To use Azure AI Speech in an application, you must create an appropriate resource in your Azure subscription. You can choose to create either of the following types of resource:

* A **Speech** resource - choose this resource type if you only plan to use Azure AI Speech, or if you want to manage access and billing for the resource separately from other services.
* An **Azure AI services** resource - choose this resource type if you plan to use Azure AI Speech in combination with other Azure AI services, and you want to manage access and billing for these services together.

## The Speech to text API

You can use Azure AI Speech to text API to perform real-time or batch transcription of audio into a text format. The audio source for transcription can be a real-time audio stream from a microphone or an audio file.

The model that is used by the Speech to text API, is based on the Universal Language Model that was trained by Microsoft. The data for the model is Microsoft-owned and deployed to Microsoft Azure. The model is optimized for two scenarios, conversational and dictation. You can also create and train your own custom models including acoustics, language, and pronunciation if the pre-built models from Microsoft do not provide what you need.

**Real-time transcription**

Real-time speech to text allows you to transcribe text in audio streams. You can use real-time transcription for presentations, demos, or any other scenario where a person is speaking.

In order for real-time transcription to work, your application will need to be listening for incoming audio from a microphone, or other audio input source such as an audio file. Your application code streams the audio to the service, which returns the transcribed text.

**Batch transcription**

Not all speech to text scenarios are real time. You might have audio recordings stored on a file share, a remote server, or even on Azure storage. You can point to audio files with a shared access signature (SAS) URI and asynchronously receive transcription results.

Batch transcription should be run in an asynchronous manner because the batch jobs are scheduled on a *best-effort basis*. Normally a job will start executing within minutes of the request but there is no estimate for when a job changes into the running state.

**The text to speech API**

The text to speech API enables you to convert text input to audible speech, which can either be played directly through a computer speaker or written to an audio file.

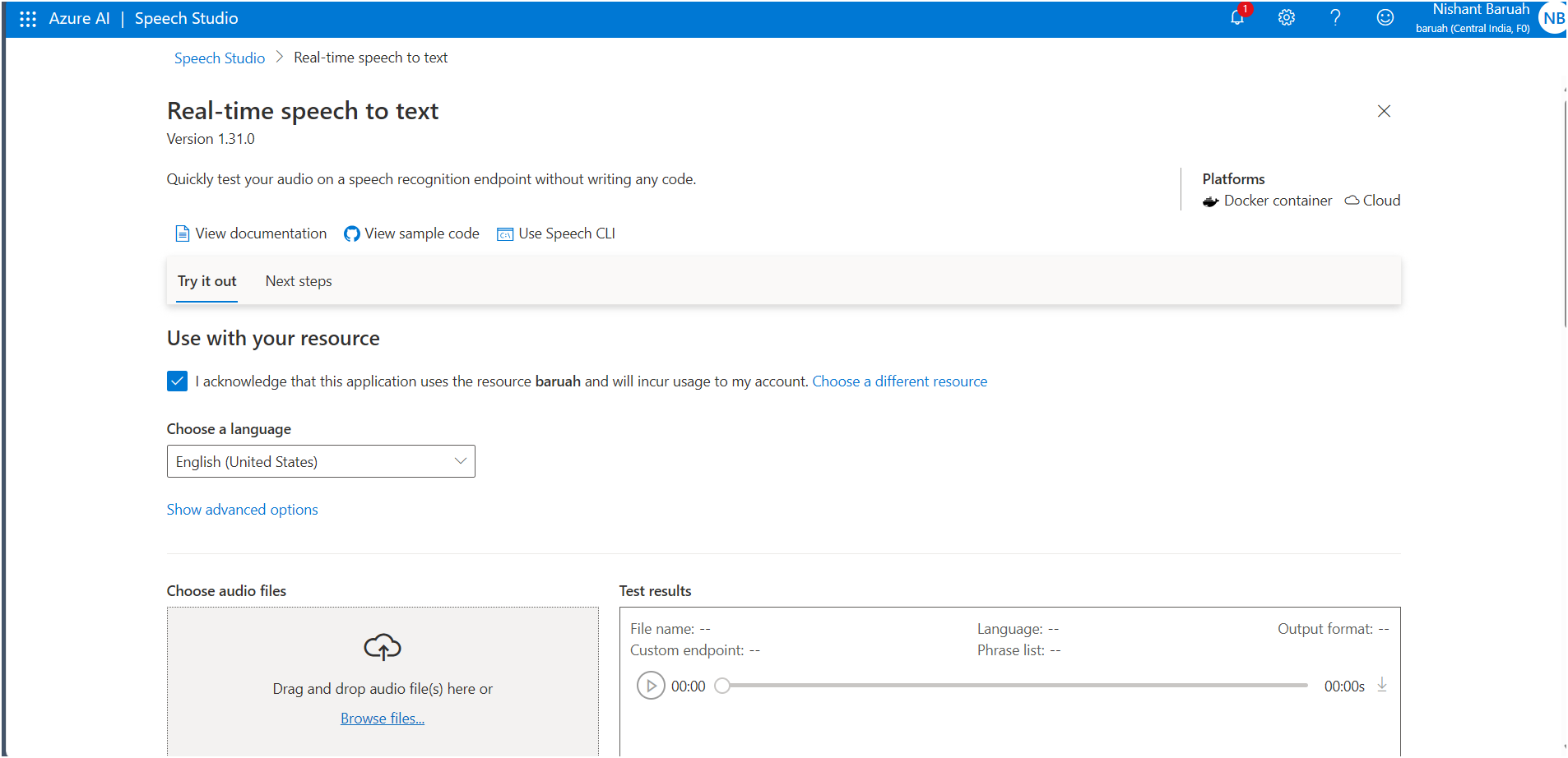
**Speech synthesis voices**

When you use the text to speech API, you can specify the voice to be used to vocalize the text. This capability offers you the flexibility to personalize your speech synthesis solution and give it a specific character.

The service includes multiple pre-defined voices with support for multiple languages and regional pronunciation, including *neural* voices that leverage *neural networks* to overcome common limitations in speech synthesis with regard to intonation, resulting in a more natural sounding voice. You can also develop custom voices and use them with the text to speech API

# Explore Speech Studio

Follow the steps in the Microsoft page



# Additional resources

Here are additional resources that provide you with an opportunity to view the languages supported by the Speech-to-text and Text-to-speech services and to see additional documentation about the Speech service.

Speech-to-text - supported languages.

<https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/language-support#speech-to-text>

Text-to-speech  - supported languages.

<https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/language-support#text-to-speech>

Service documentation about the Speech service

<https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/>

## Translate text and speech

# What is literal and semantic translation?

* The video lecture is about performing text and speech translation using cognitive services on Microsoft Azure.
* It discusses the problem of language barriers in collaboration between people from different cultures and geographic locations.
* Automated translation, also known as machine translation, is being used to solve this problem.
* Early attempts at machine translation used literal translations, where each word is translated to the corresponding word in the target language.
* However, literal translation can have issues, such as not having an equivalent word in the target language or changing the meaning of the phrase.
* A translation service should take into account the semantic context to provide a more accurate translation.
* Semantics, the study of relationships between words and how we construct meaning, is important for understanding and translating languages.
* Artificial intelligence systems need to understand not only the words but also the semantic context in which they're used.
* Grammar rules, formality, and colloquialisms also need to be considered in translation.

What are some challenges of using literal translation in machine translation?

Some challenges of using literal translation in machine translation include:

* Lack of equivalent words: Literal translation may not have an equivalent word in the target language, leading to inaccurate or nonsensical translations.
* Changing the meaning: Literal translation can change the meaning of a phrase or fail to capture the intended context, resulting in incorrect translations.
* Idiomatic expressions: Literal translation may not account for idiomatic expressions or colloquialisms, leading to translations that sound unnatural or confusing.
* Grammar and syntax differences: Literal translation may not consider the grammar and syntax differences between languages, resulting in grammatically incorrect translations.
* Cultural nuances: Literal translation may not capture cultural nuances or specific cultural references, leading to translations that are not culturally appropriate or understandable.

It is important for translation services to go beyond literal translation and consider the semantic context, grammar rules, idiomatic expressions, and cultural nuances to provide more accurate and meaningful translations.

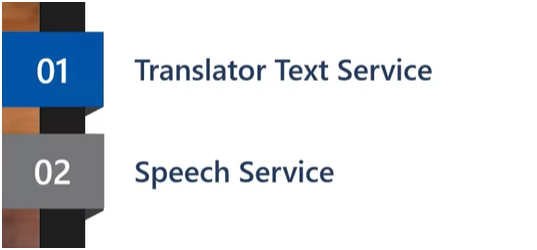
Can you explain the role of semantics in understanding and translating languages?

Semantics plays a crucial role in understanding and translating languages. It refers to the study of the relationships between words and how we construct meaning. Here's how semantics impacts language understanding and translation:

1. Contextual understanding: Semantics helps us understand the meaning of words and phrases in different contexts. Words can have multiple meanings, and semantics helps determine the intended meaning based on the surrounding words and the overall context of the sentence or conversation.
2. Ambiguity resolution: Languages often contain ambiguous words or phrases that can be interpreted in different ways. Semantics helps in resolving these ambiguities by considering the context and the intended meaning. It allows us to choose the most appropriate interpretation based on the semantic context.
3. Idiomatic expressions: Semantics helps in understanding and translating idiomatic expressions, which are phrases that have a figurative meaning different from the literal interpretation. Without considering the semantic context, translating idioms can lead to confusion or loss of meaning. Semantics helps capture the intended figurative meaning and provide accurate translations.
4. Cultural nuances: Different languages and cultures have their own unique expressions, proverbs, and cultural references. Semantics helps in understanding these cultural nuances and translating them appropriately. It considers the cultural context to ensure that translations are culturally accurate and meaningful.
5. Machine translation accuracy: In machine translation, semantics is essential for accurate and meaningful translations. By incorporating semantic understanding, translation systems can go beyond literal translations and consider the intended meaning, idiomatic expressions, and cultural nuances. This leads to more accurate and contextually appropriate translations.

Overall, semantics provides the foundation for understanding the meaning of words and phrases in different contexts, resolving ambiguities, capturing idiomatic expressions, considering cultural nuances, and improving the accuracy of language translation.

# Get Started with Translation in Azure



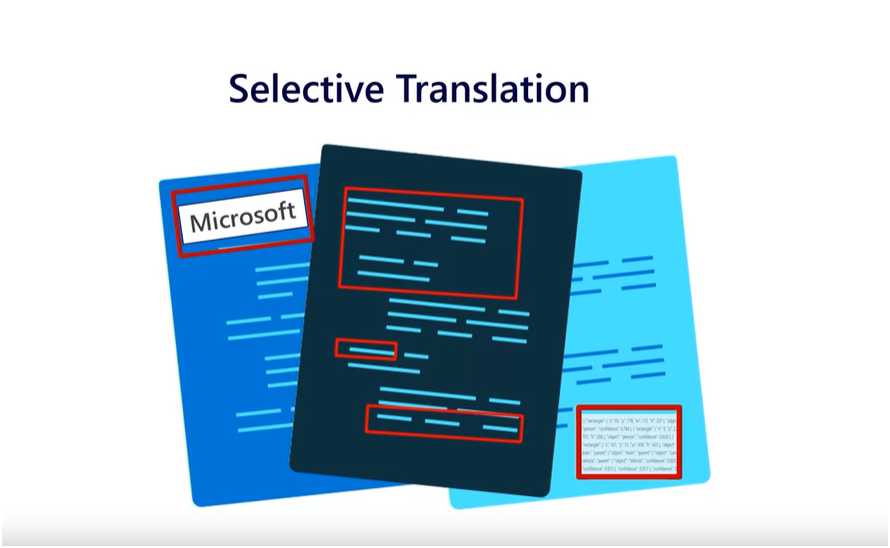




Profanity filtering is a feature that allows you to control the translation of profane or offensive language when using the Translator Text service in Microsoft Azure. With profanity filtering, you have two options:

1. Marking the translated text as profane: This means that if the input text contains profanity, the translated text will also include the profanity. It is useful when you want to maintain the original context of the text.
2. Omitting the profanity in the results: This option filters out any profanity from the translated text. It is helpful when you want to ensure that the translated text is free from offensive language.

By using profanity filtering, you can customize the translation results based on your specific requirements and the cultural context of the target language.



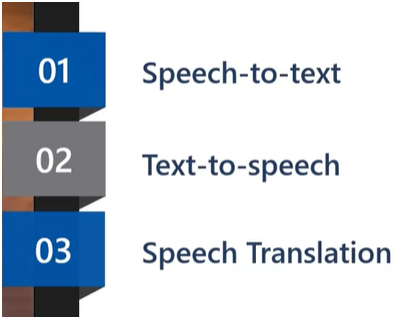
Selective translation is a feature offered by the Translator Text service in Microsoft Azure. It allows you to tag specific content that you do not want to be translated. For example, you can tag code snippets, brand names, or phrases that may not make sense when localized.

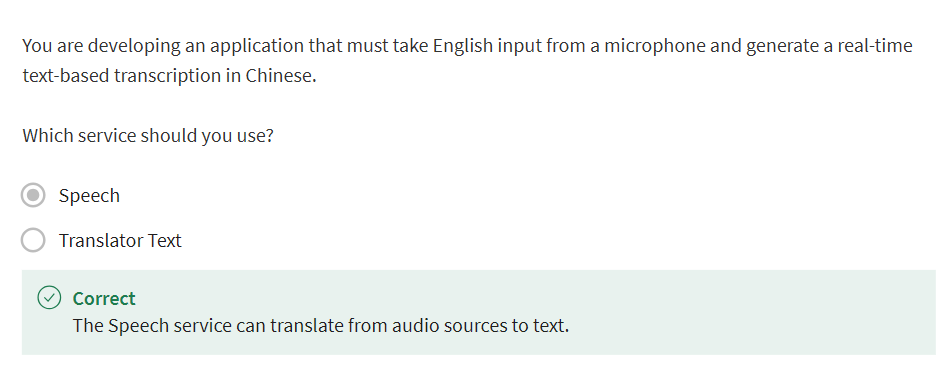
By using selective translation, you can ensure that certain parts of your text remain in their original language while translating the rest. This feature provides more control over the translation process and allows you to preserve the intended meaning and context of the tagged content.

To use selective translation, you can specify the content that should not be translated by applying appropriate tags or markers. This way, when you use the Translator Text API, the tagged content will be excluded from the translation process, and only the remaining text will be translated.

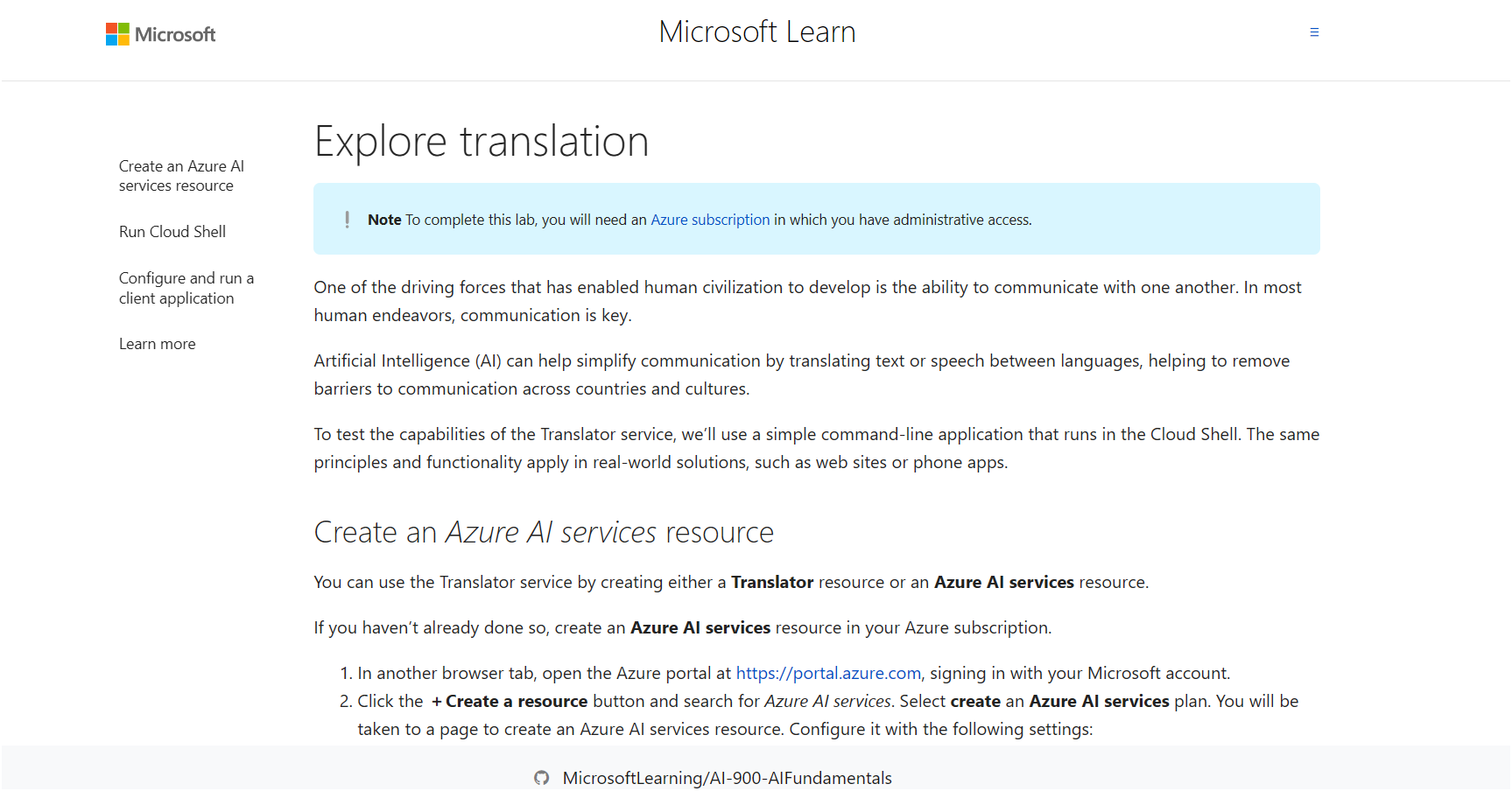
Selective translation is particularly useful when you have specific requirements for certain parts of your text that should not be altered during the translation process.

Speech Service

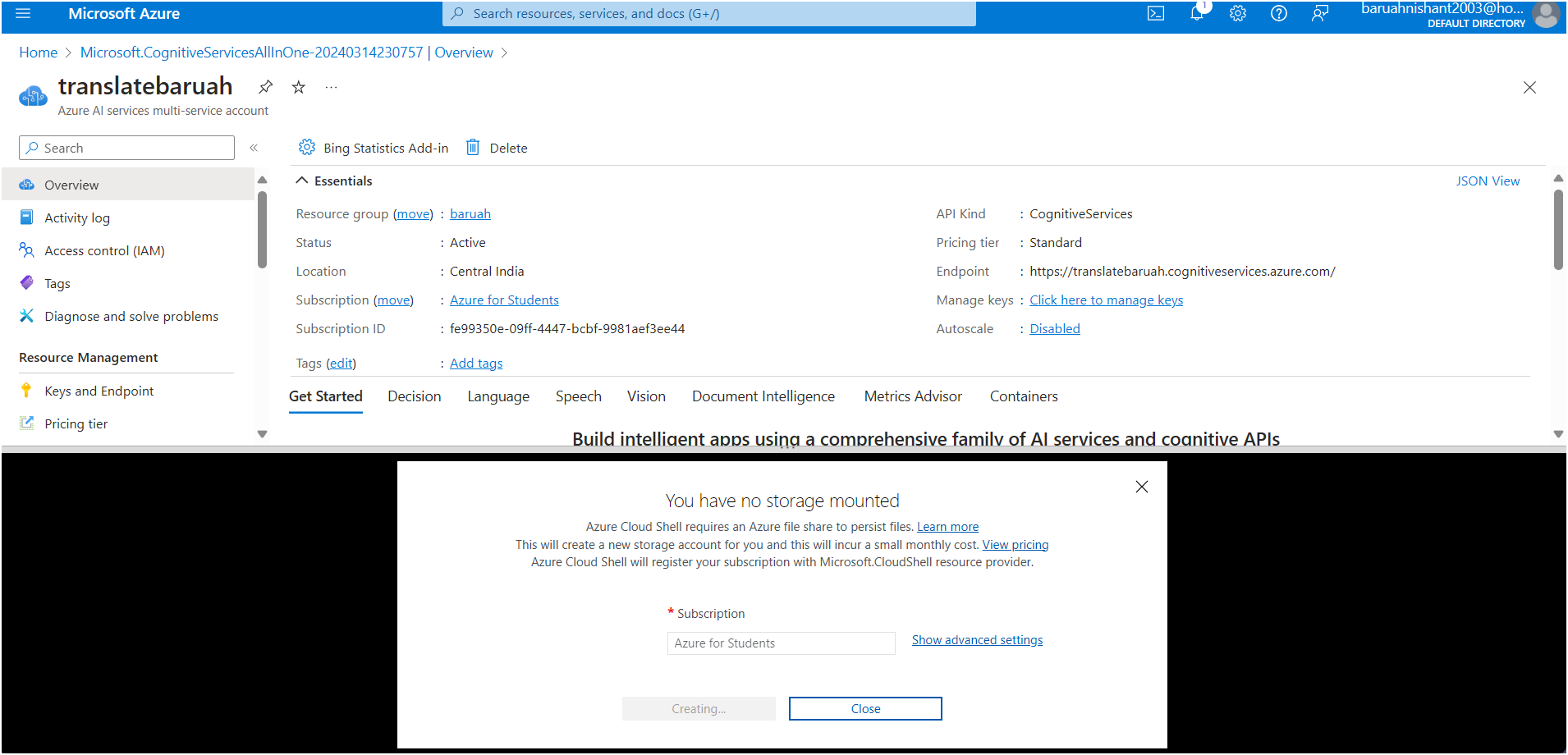


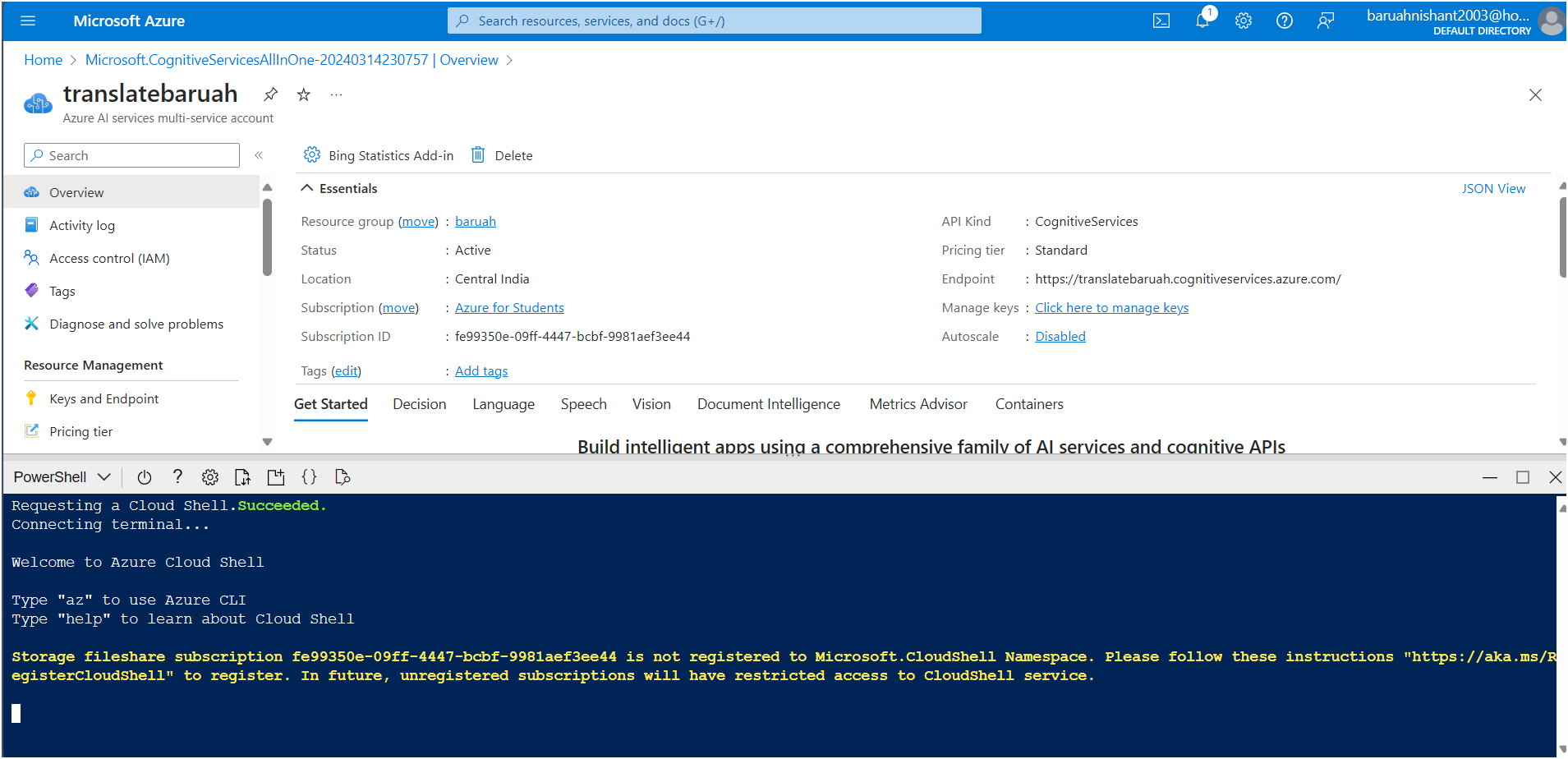


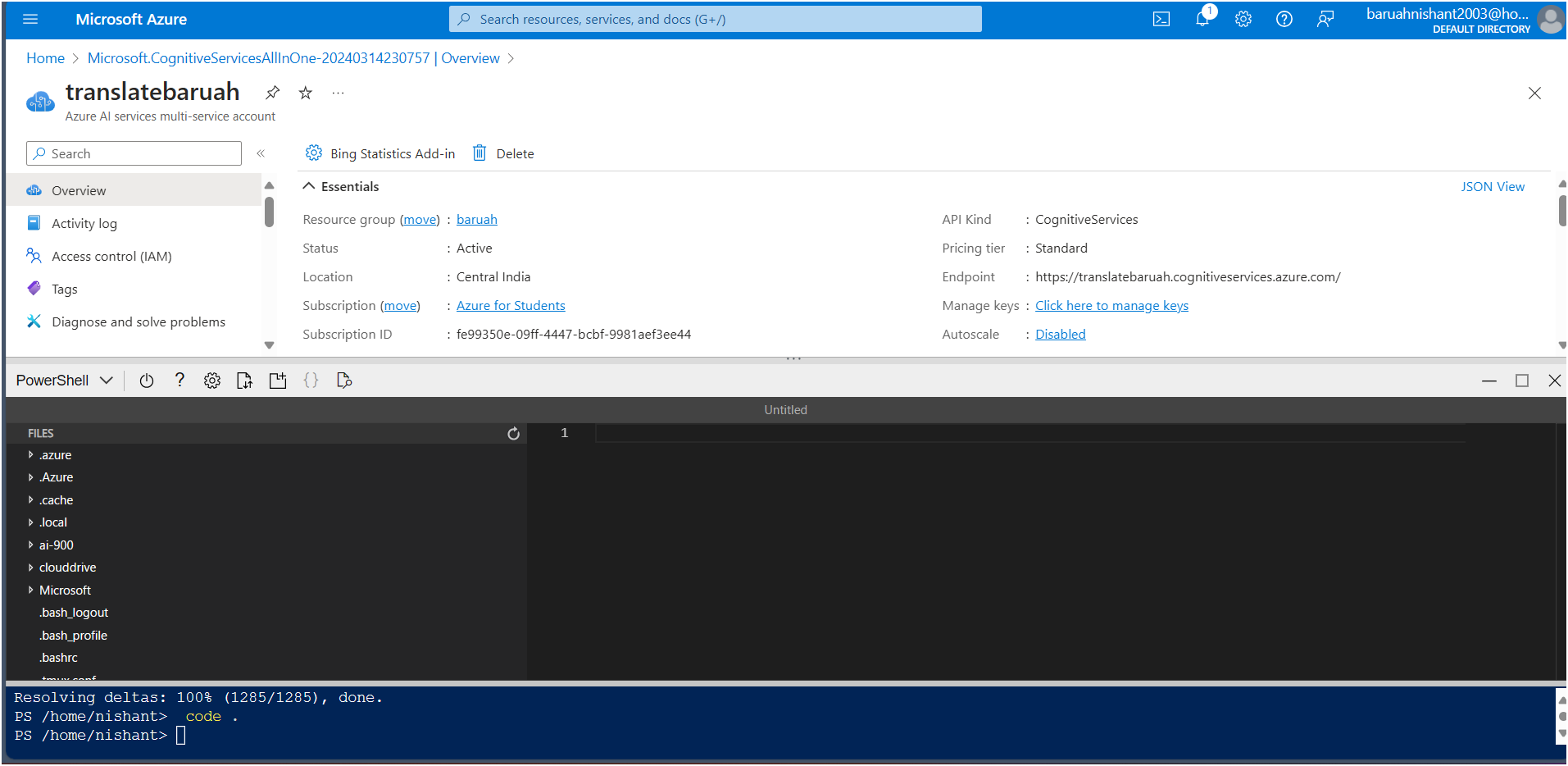
# Explore translation

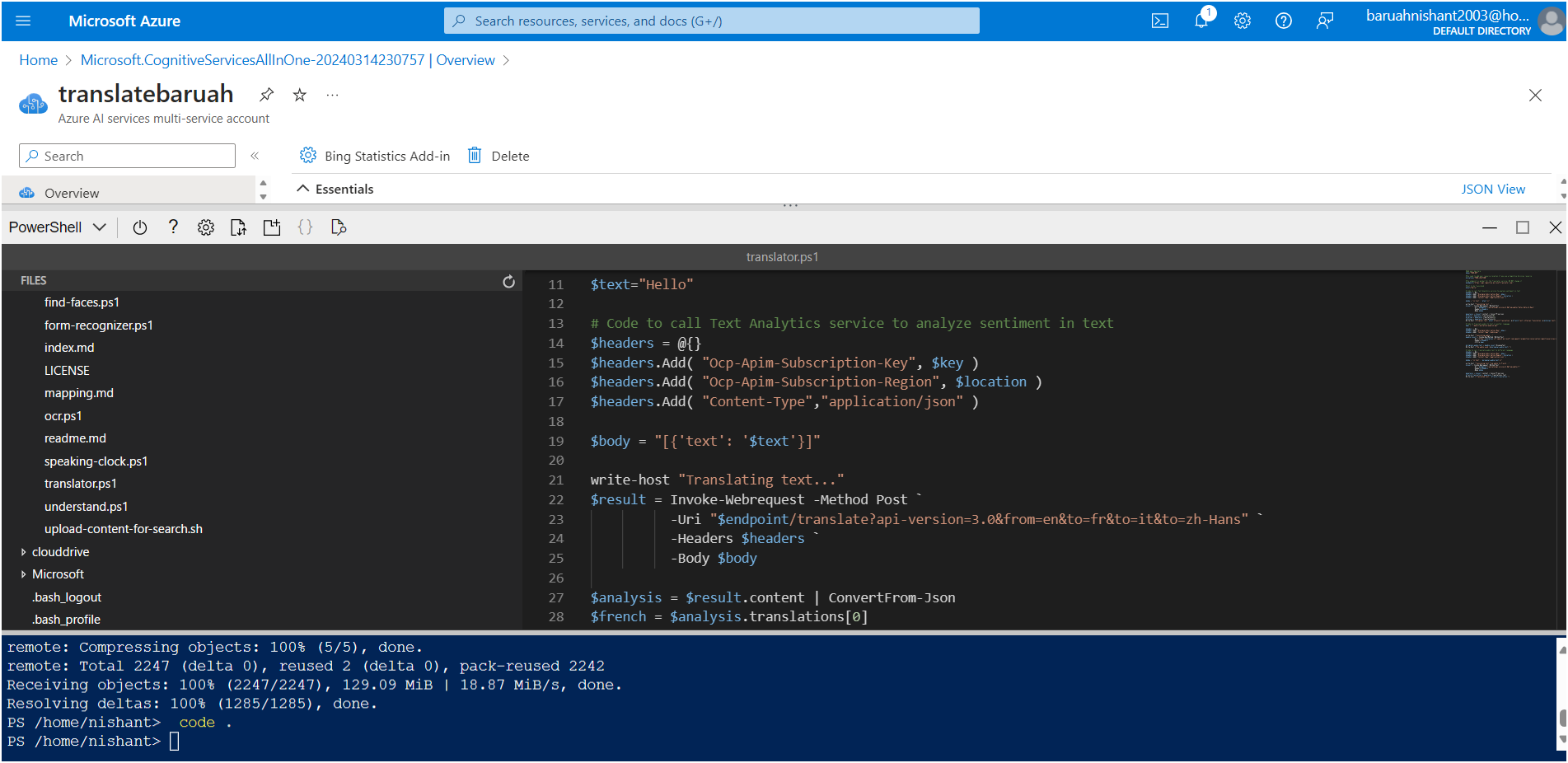


Follow the instruction to complete **Exercise - Explore translation**

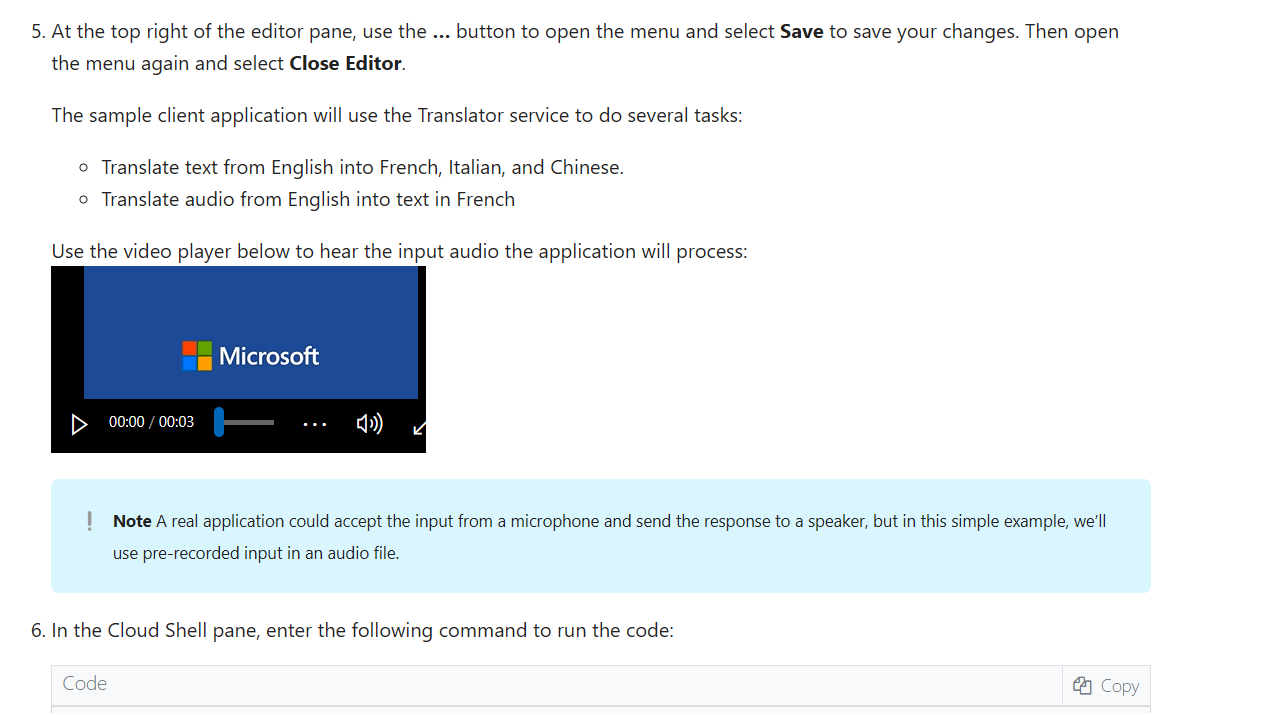
****

****

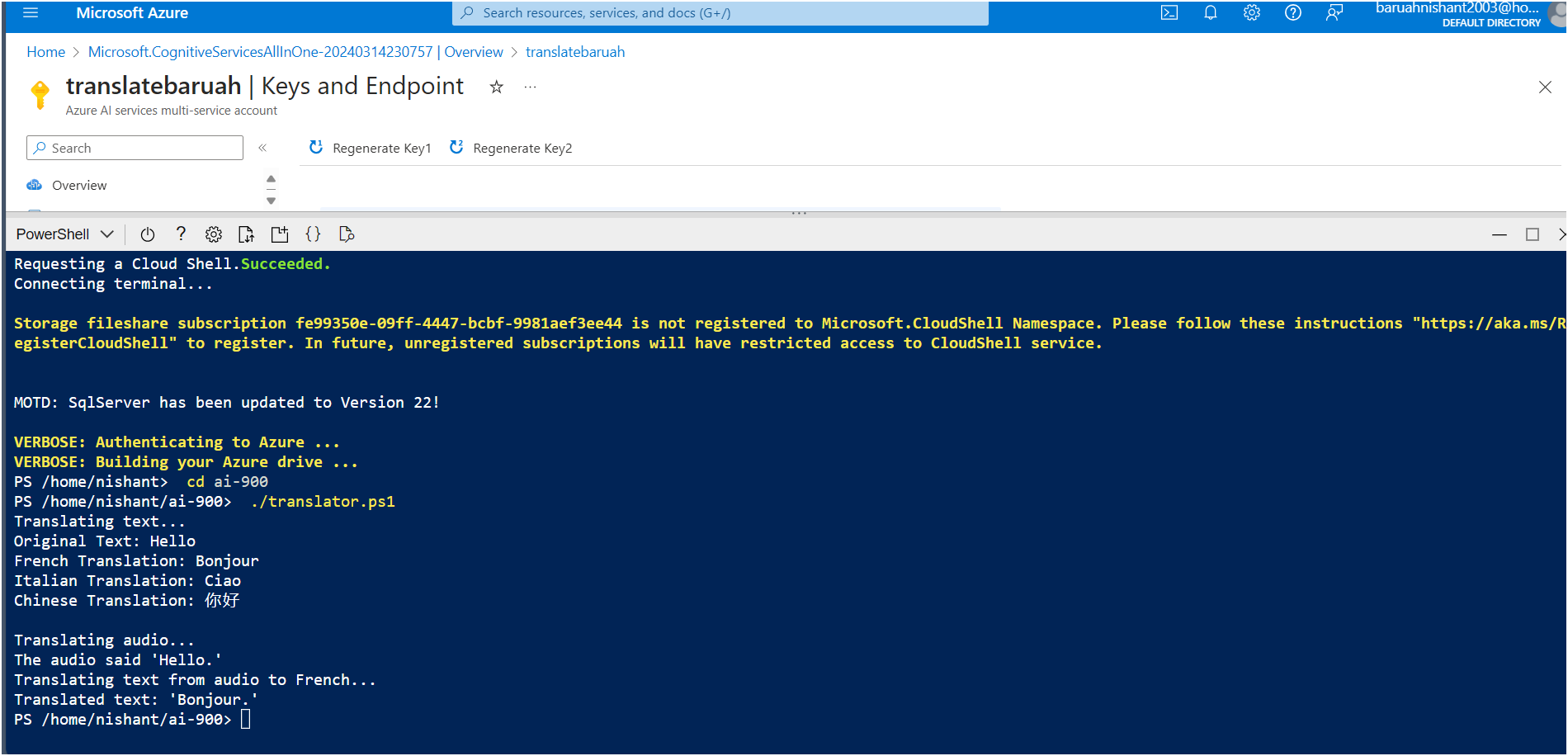
****

****

**Translatting the audio tp French, Italian and Chinese**

****

**Result:**

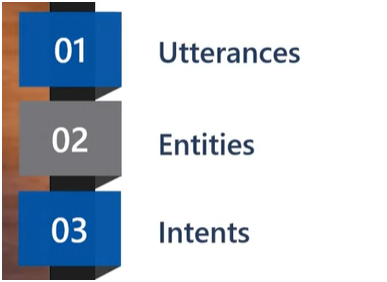
****

## **Learn more**

This simple app shows only some of the capabilities of the Translator service. To learn more about what you can do with this service, see the [Translator page](https://docs.microsoft.com/azure/cognitive-services/translator/translator-overview).

## Create a Language Model with Language Understanding

# What is Language Understanding?



In the context of natural language processing, an utterance refers to a piece of text or speech that a user might say or input. It is an example of language that your application needs to interpret and understand. Utterances are used to train language models and identify the intent and entities within the user's input. For example, in a home automation system, utterances like "switch the fan on" or "turn on the light" would be examples of utterances that the system needs to interpret.

Can you explain the purpose of intents in language understanding applications?How can entities be used to extract specific information from user utterances?What is the significance of the non intent in a language understanding application?

An entity, in the context of natural language processing, refers to a word or phrase within an utterance that represents a specific object or concept. Entities are used to extract relevant information from the user's input. For example, in the utterance "Book a flight from New York to London," the entities would be "New York" and "London," representing the departure and destination locations. Entities help in understanding the user's intent and enable the application to provide more accurate and personalized responses.

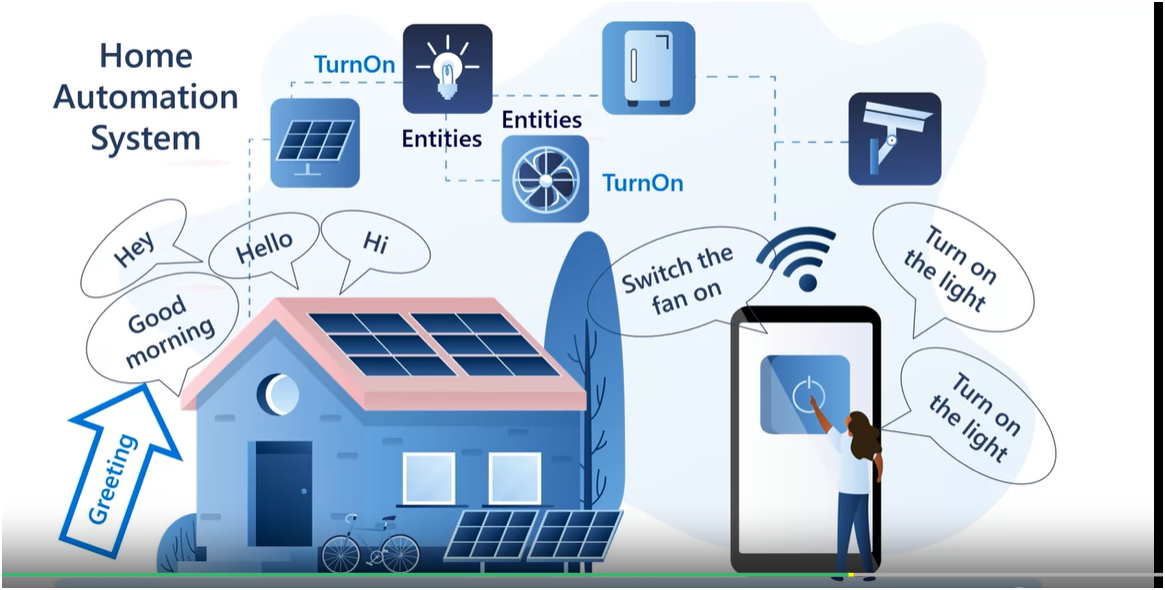
An intent, in the context of natural language processing, represents the purpose or goal expressed in a user's utterance. It is a way of categorizing the tasks or actions that the user wants to perform. For example, if a user says "Book a flight from New York to London," the intent would be to book a flight. Intents help in understanding the user's intention and enable the application to provide appropriate responses or take the necessary actions. In language understanding applications, intents are defined and trained using a set of sample utterances that are associated with each intent.





Intent-



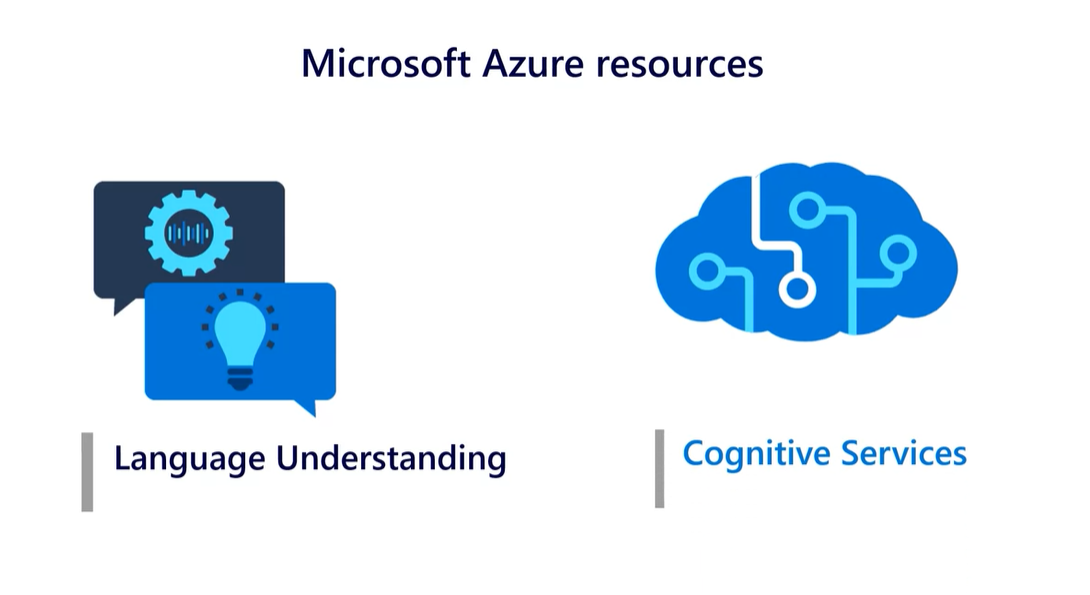


Commands are utterances

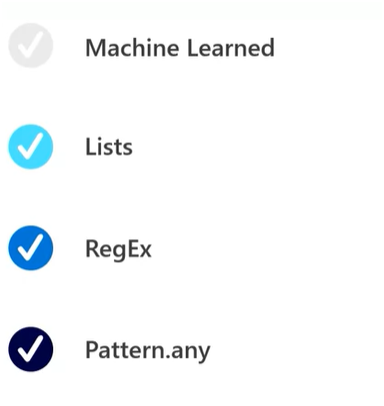
Devices are entity

‘Turn On’ is intent

# Getting Started with Language Understanding

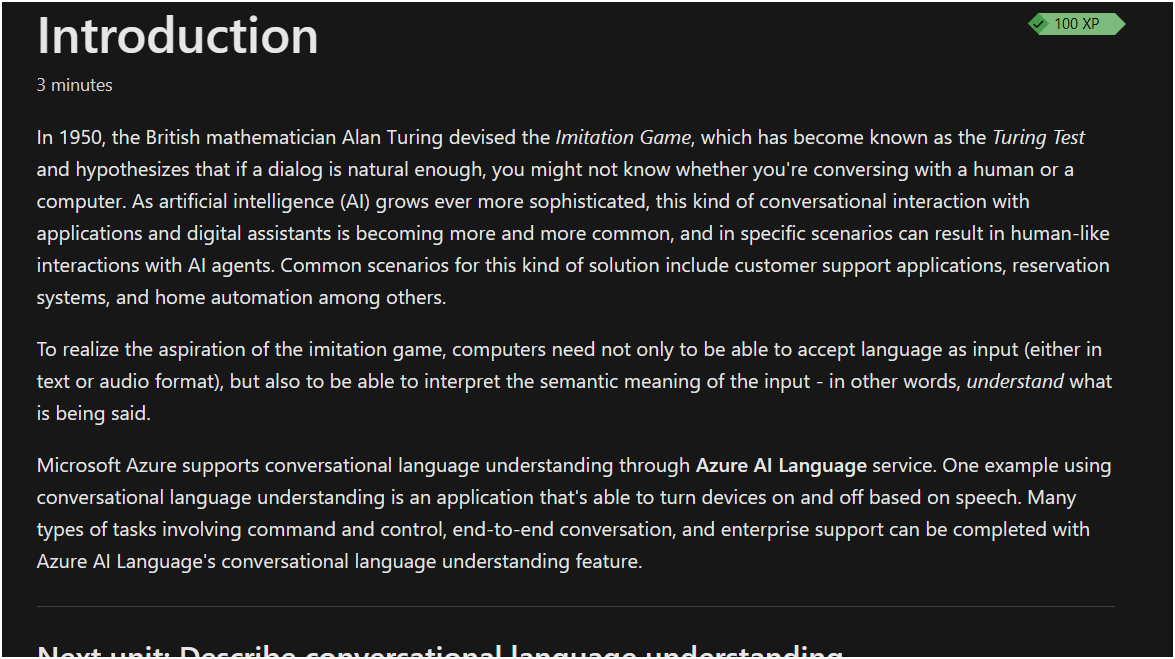


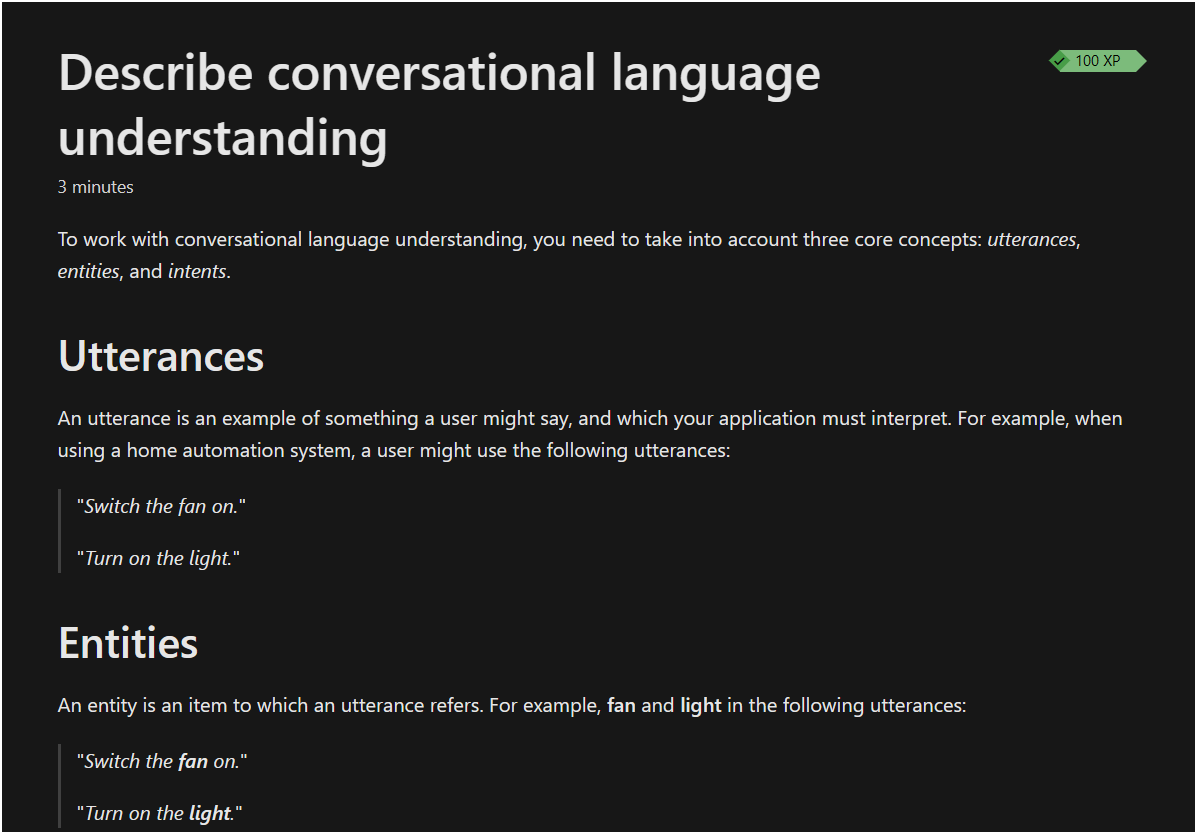
Four types of entities:

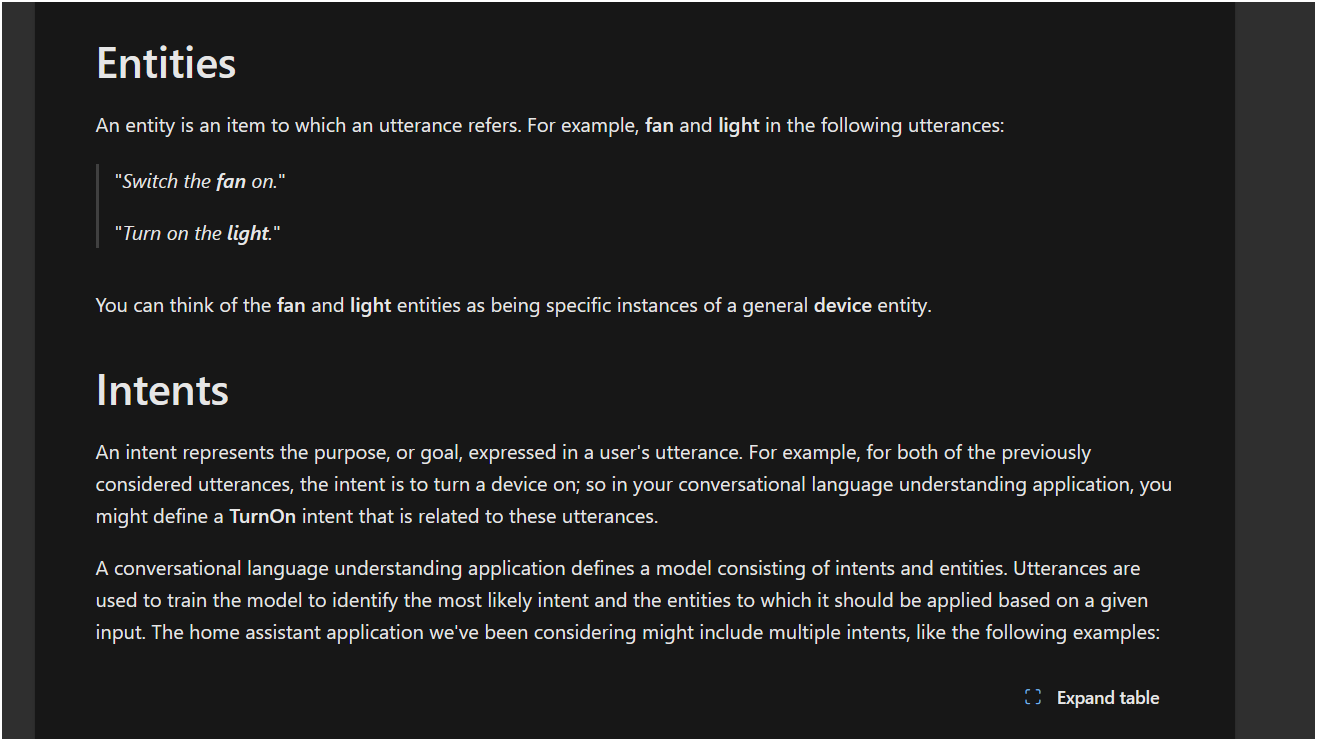




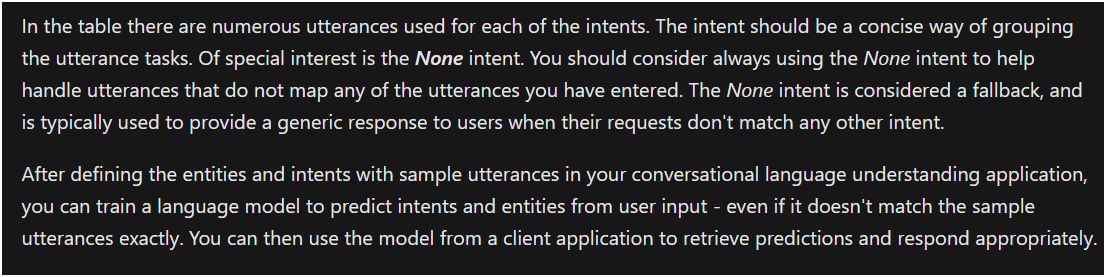
Follow the steps in Microsoft learn docs website:

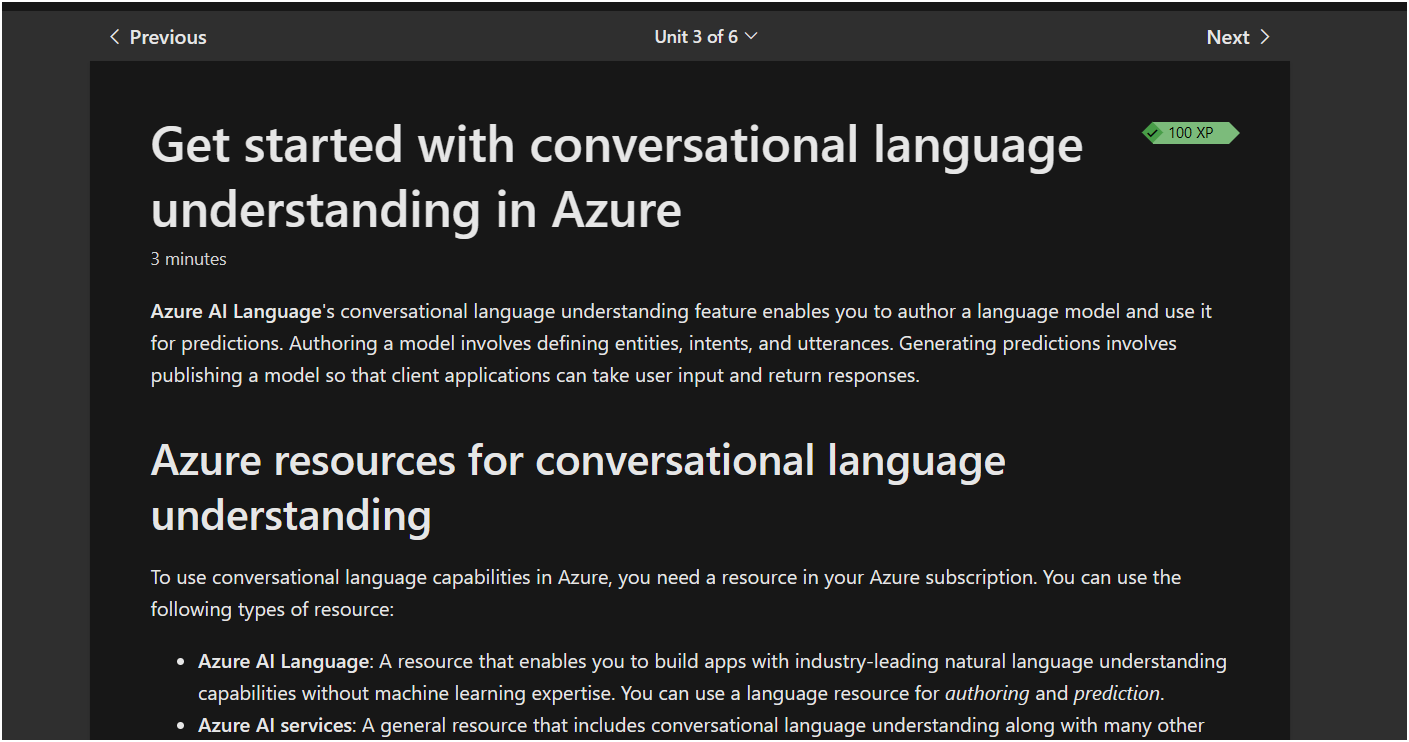




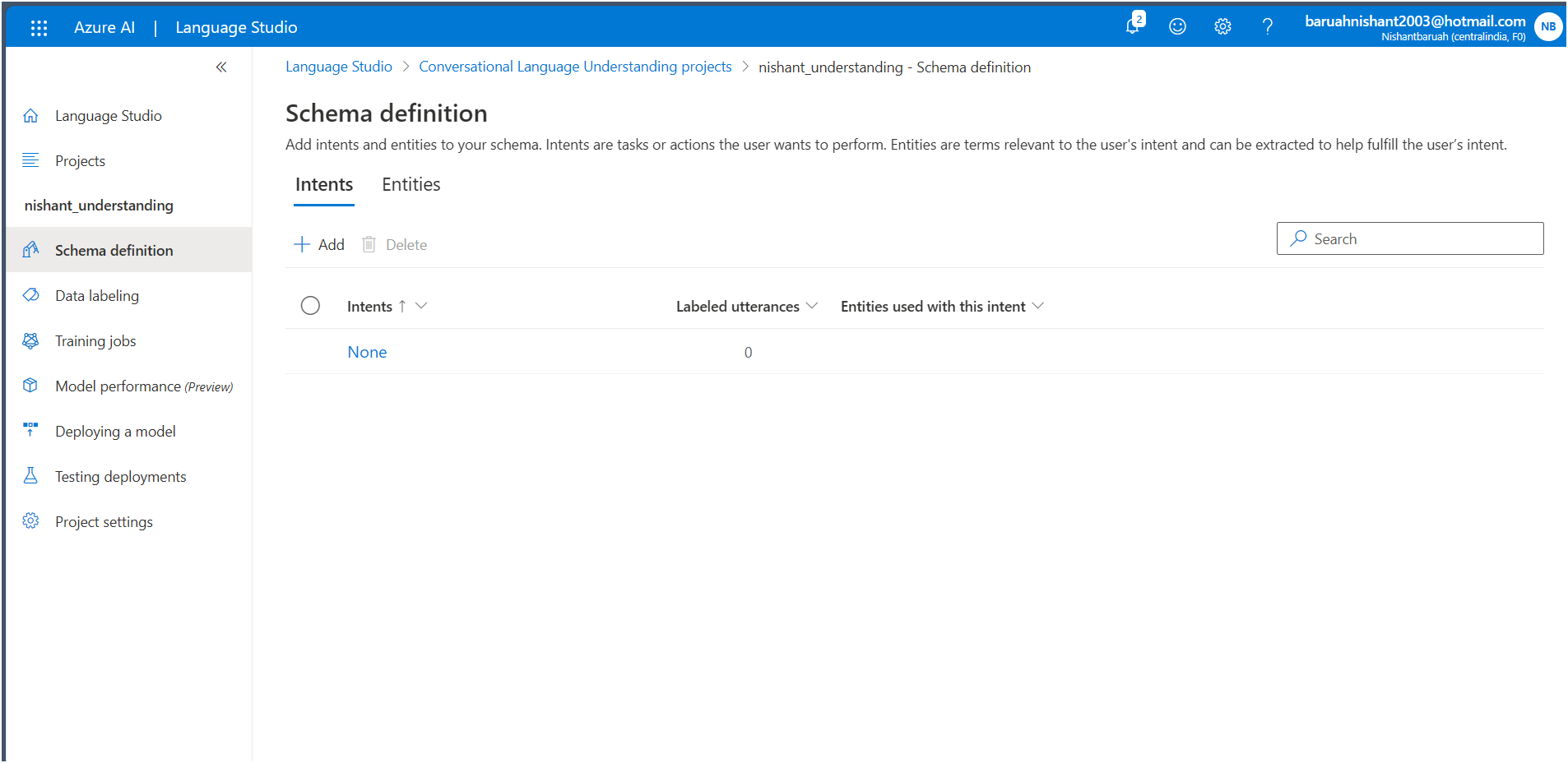




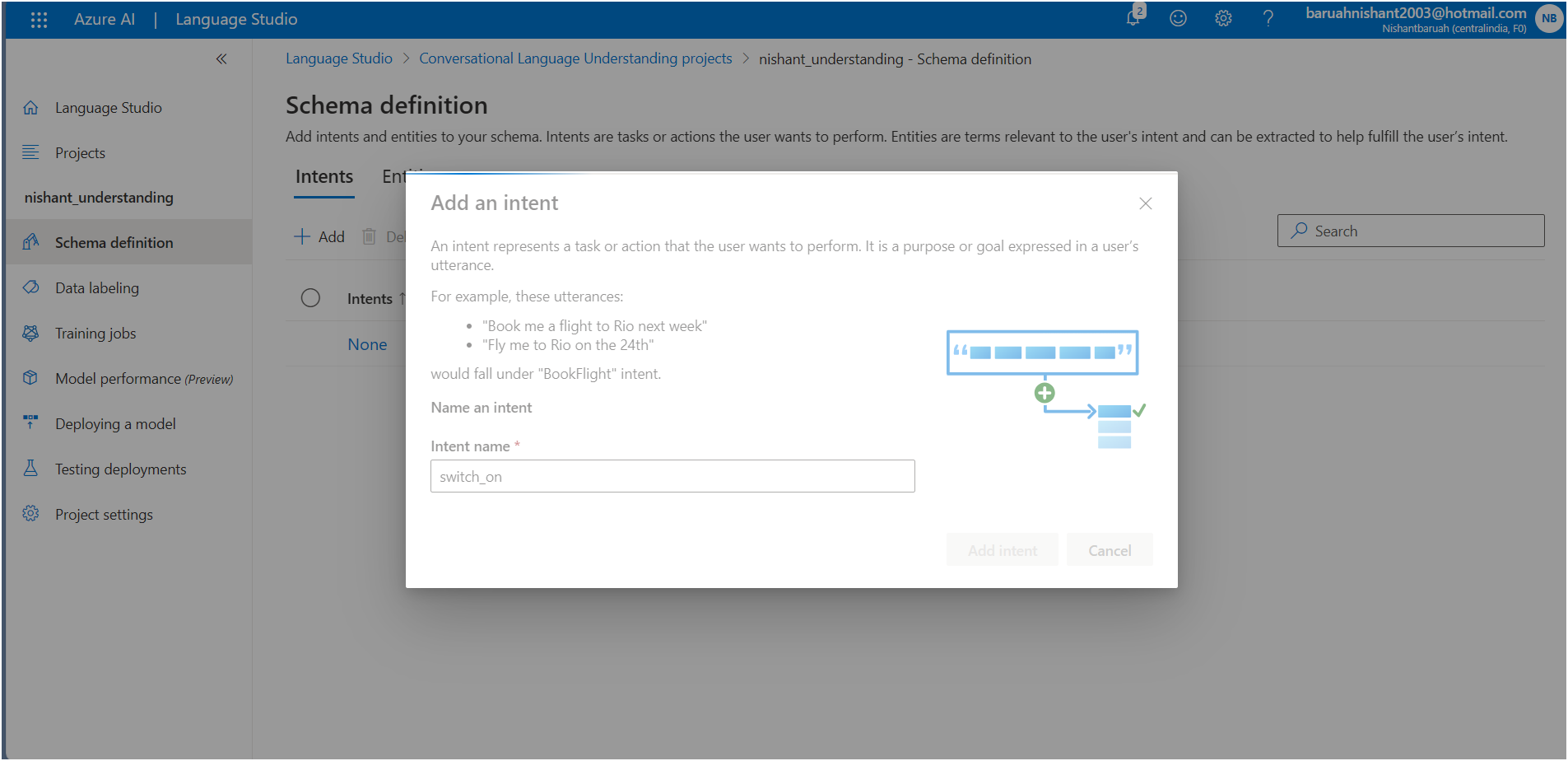


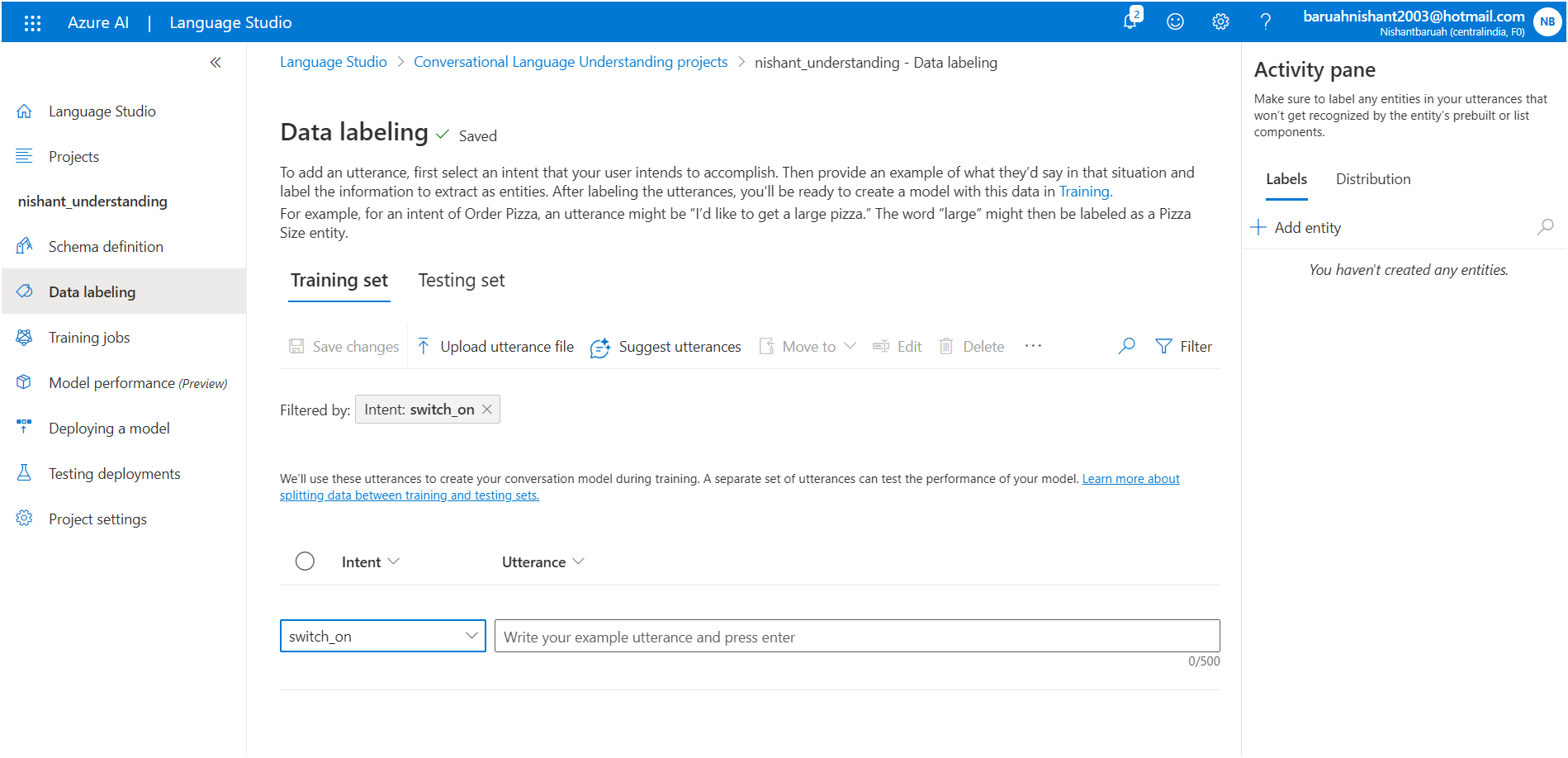


Results:

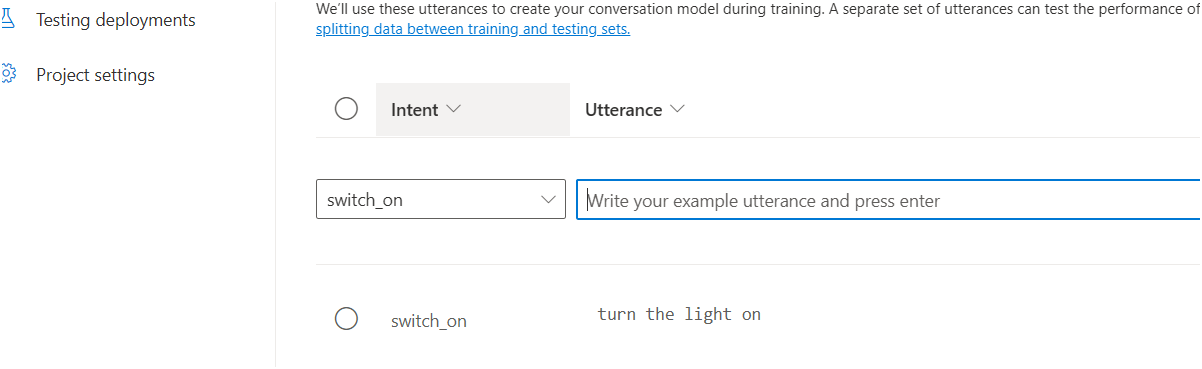


Adding intent-

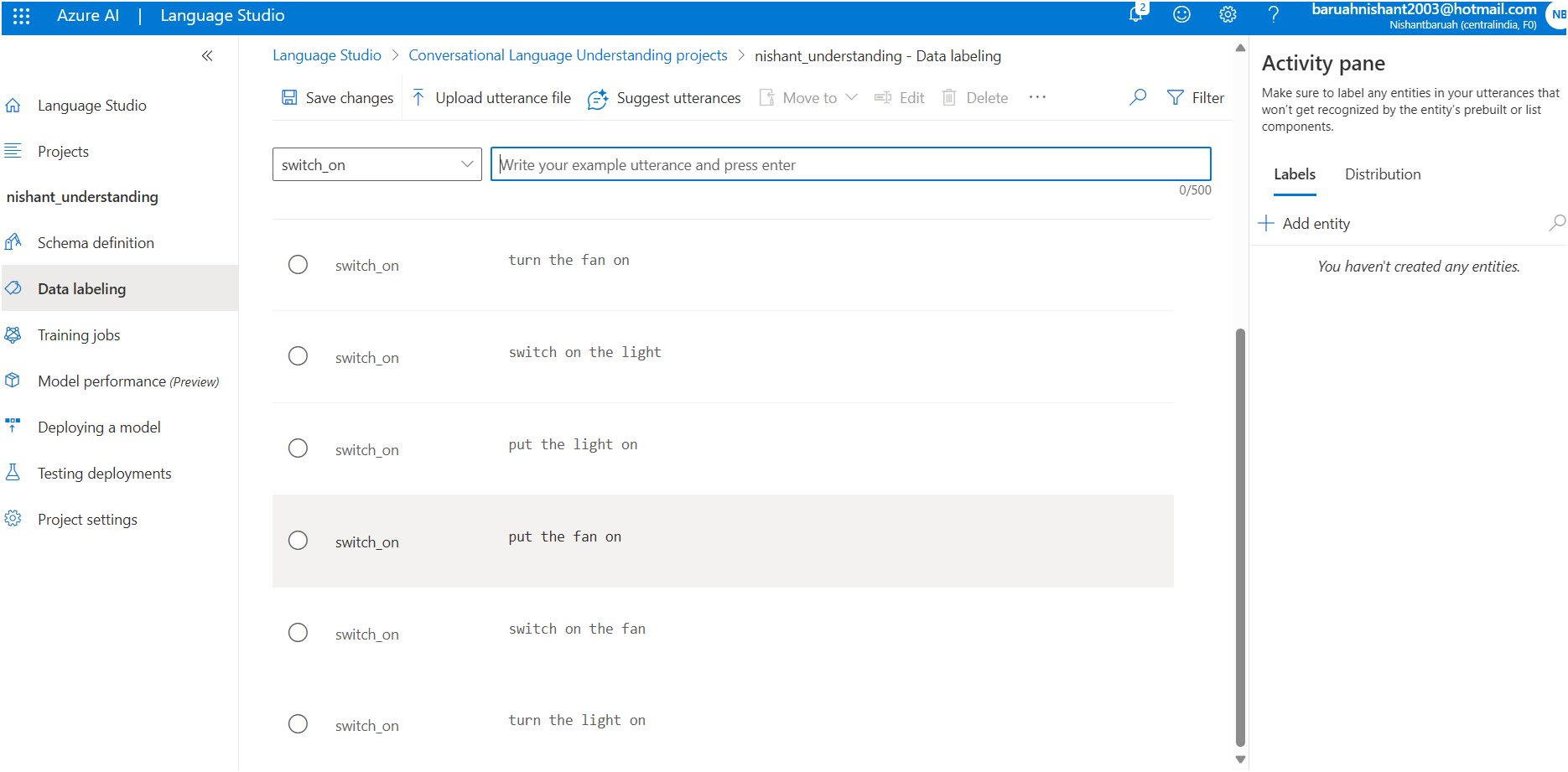


****

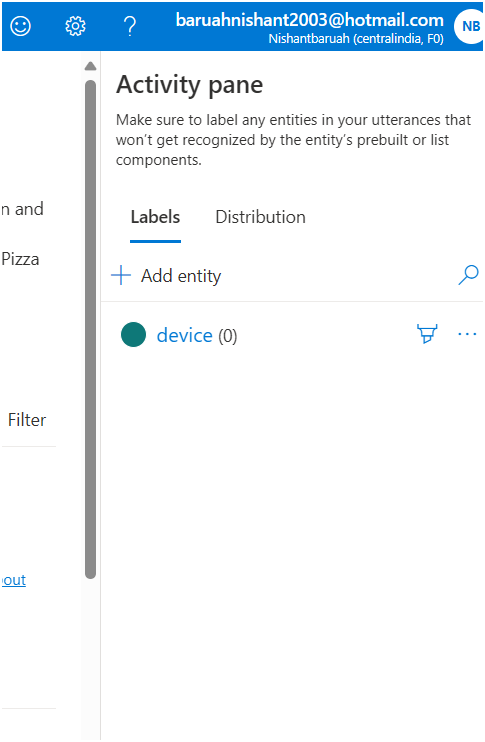
**Intent and utterance added-**

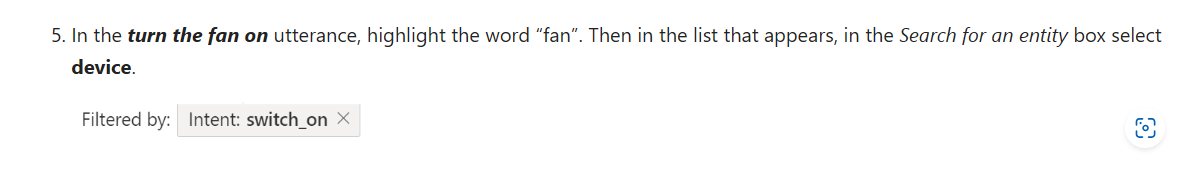
****

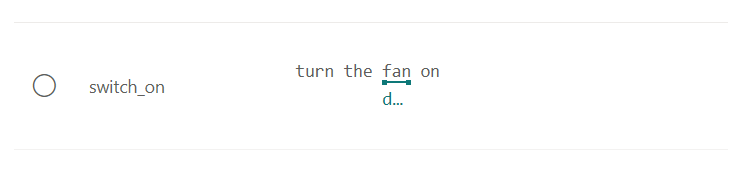
**More utterances for the same intent**

****

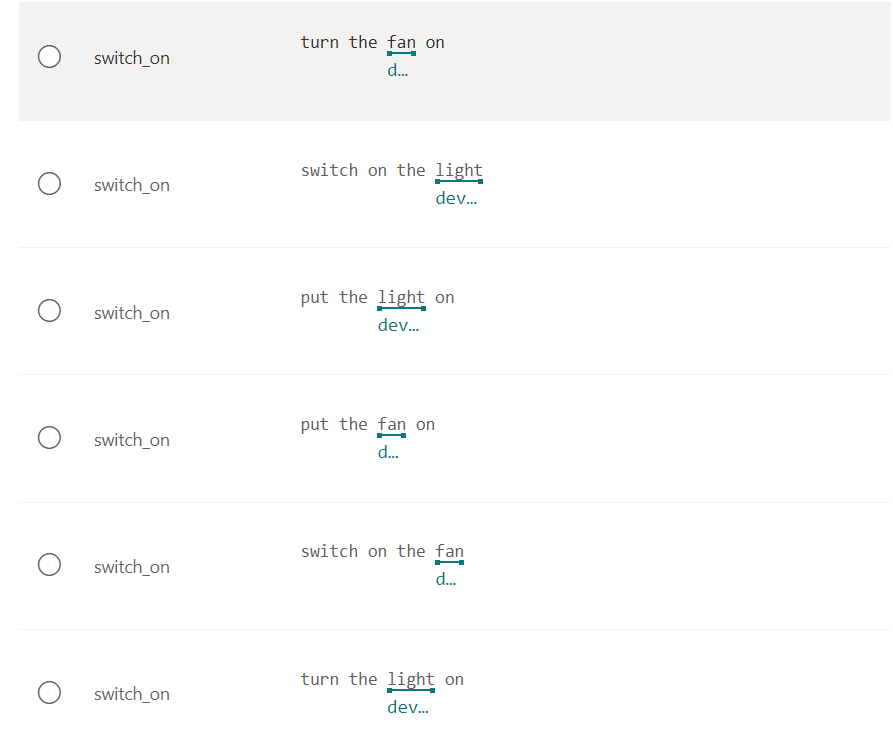
**Entity added-**

****

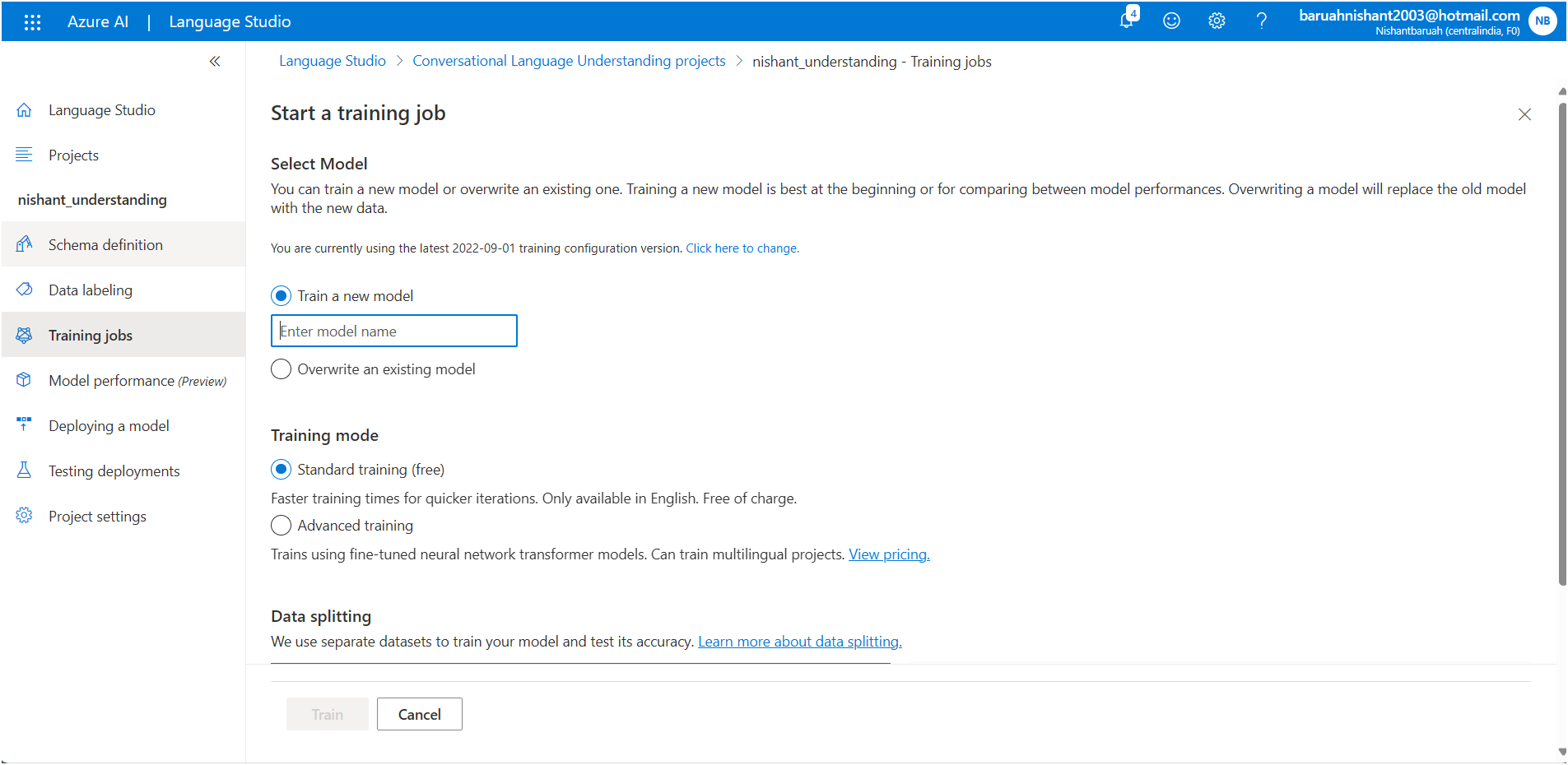
****

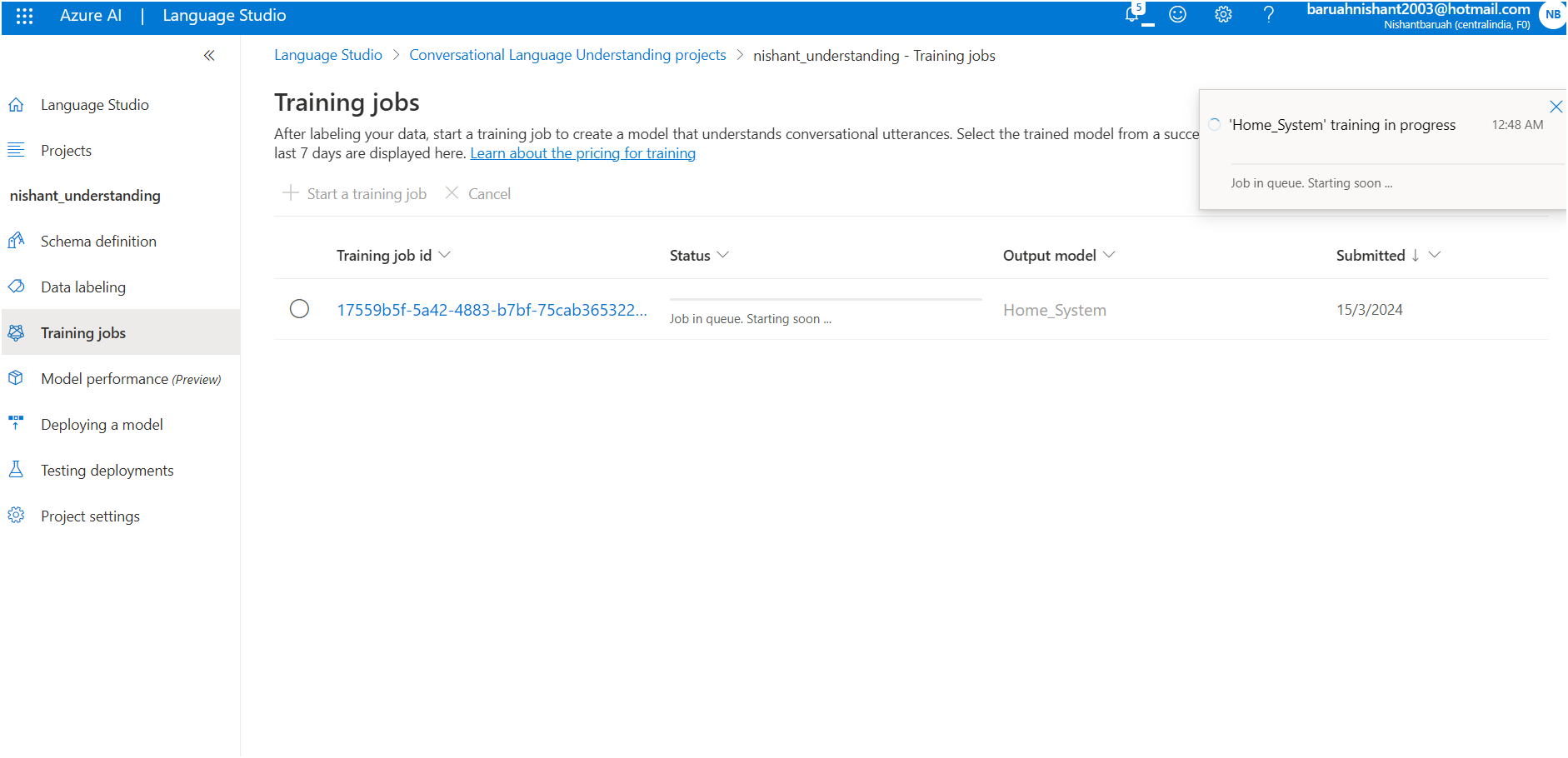
****

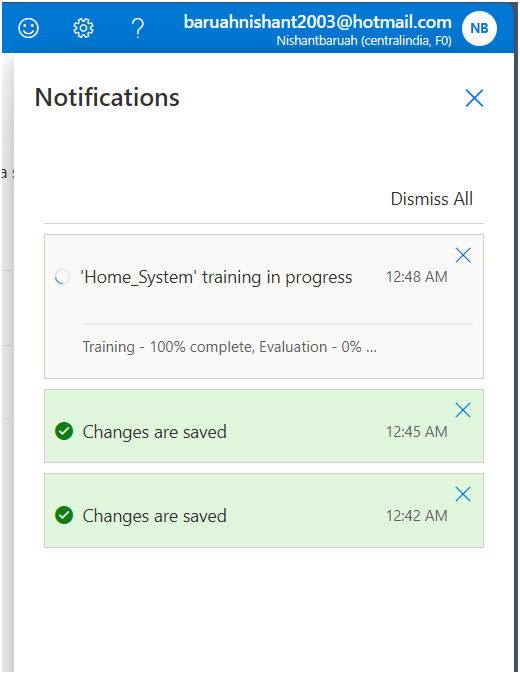
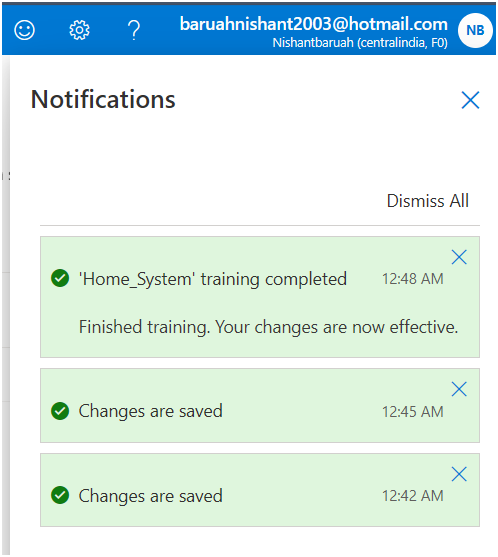
Label the rest of the fan or light utterances with the **device** entity

****

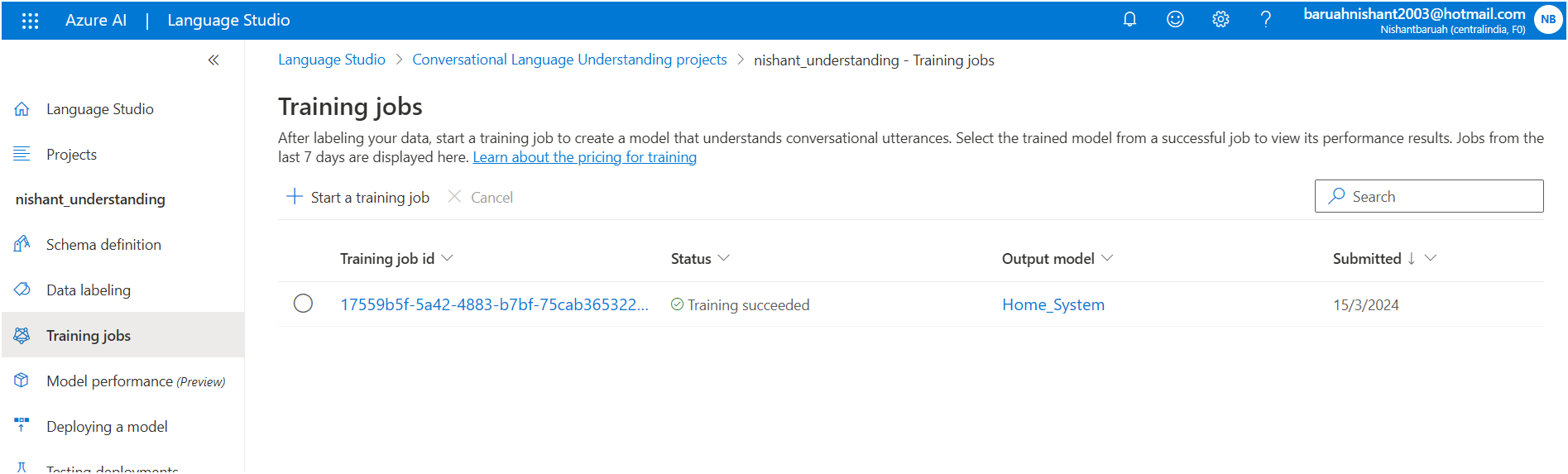
Training the jobs:



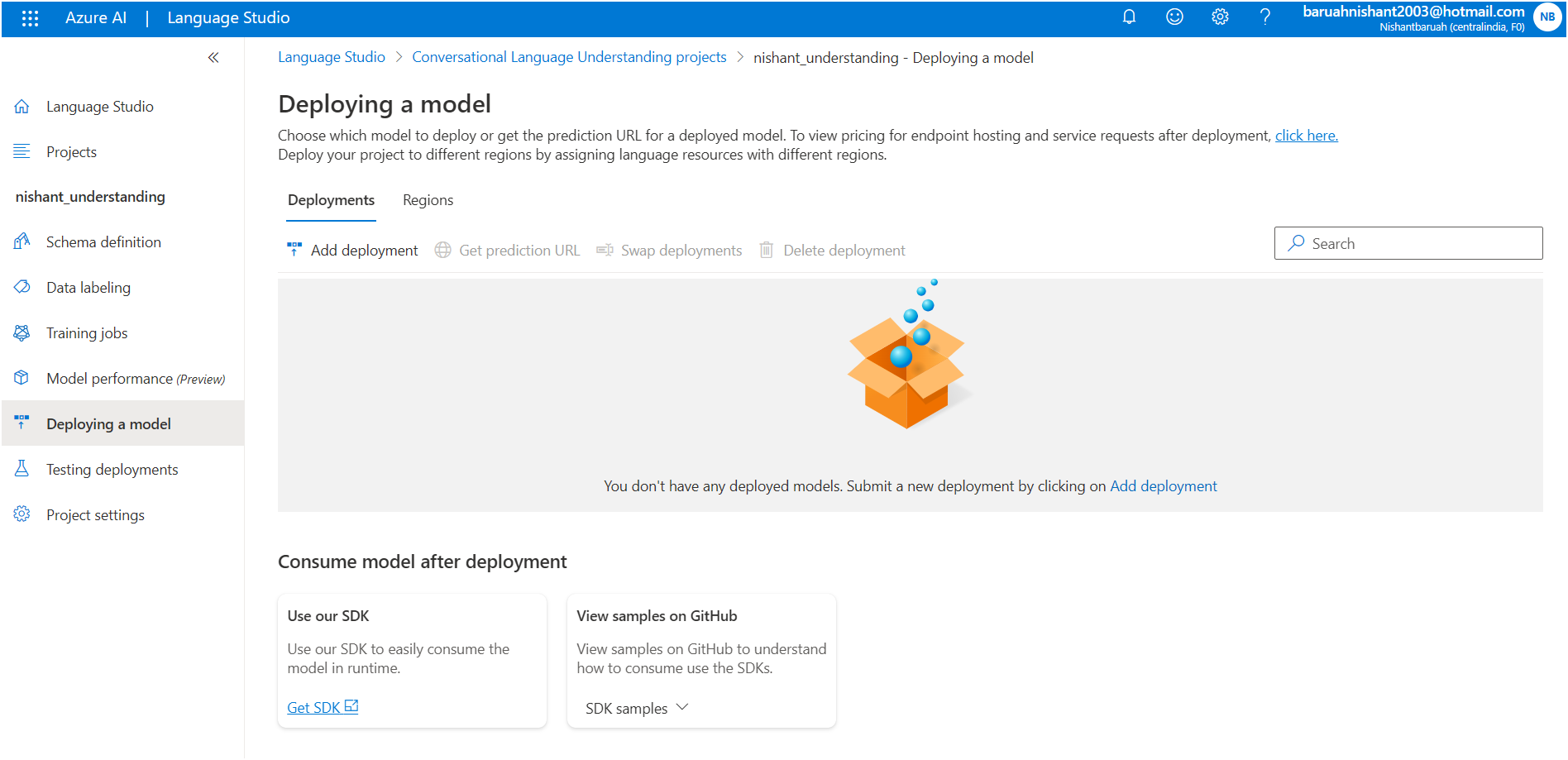


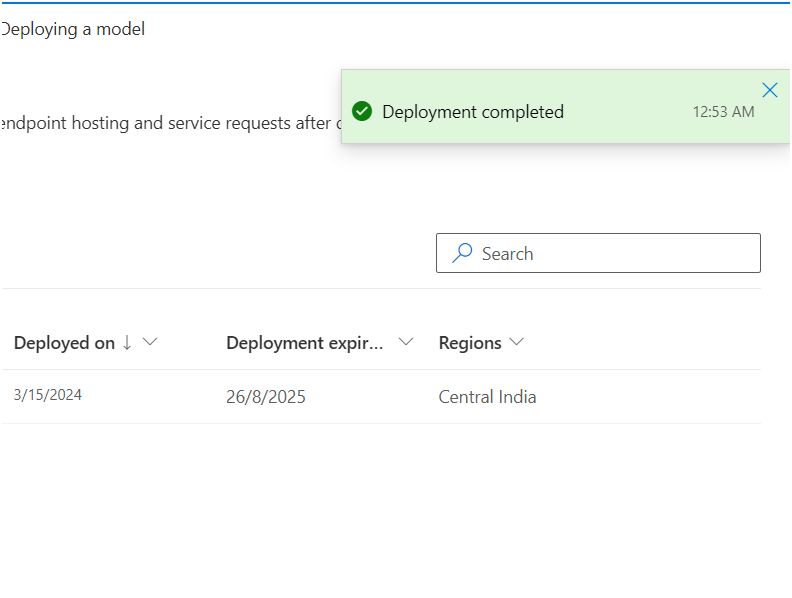
Model trained:



Now deploying the model:



Deployed



# Summary

Completed100 XP

* 1 minute

You can use conversational language understanding to build a model that predicts intents and entities from natural language utterances. A client application can then use this trained model to respond to natural language user input.

You can find out more about conversational language understanding in the [Azure AI Language documentation](https://learn.microsoft.com/en-us/azure/ai-services/language-service/conversational-language-understanding/overview).

## Clean-up

It's a good idea at the end of a project to identify whether you still need the resources you created. Resources left running can cost you money.

If you are continuing on to other modules in this learning path you can keep your resources for use in other labs.

If you have finished learning, you can delete the resource group or individual resources from your Azure subscription:

1. In the [Azure portal](https://portal.azure.com/), in the **Resource groups** page, open the resource group you specified when creating your resource.
2. Click **Delete resource group**, type the resource group name to confirm you want to delete it, and select **Delete**. You can also choose to delete individual resources by selecting the resource(s), clicking on the three dots to see more options, and clicking **Delete**.

## Build a Bot with QnA Maker and Azure Bot Service

# What is Conversational AI?

Conversational AI refers to the technology that enables computers to interact with humans in a conversational manner, just like a real conversation. It involves the use of artificial intelligence (AI) techniques to understand and respond to human language, whether it's spoken or written. Conversational AI systems, often called chatbots or virtual assistants, can understand user queries, provide relevant information, answer questions, and even engage in more complex conversations. These systems use natural language processing (NLP) algorithms to interpret and understand the meaning behind human language, allowing them to generate appropriate responses. Conversational AI is used in various applications, such as customer support, virtual assistants, voice-controlled devices, and more.

# Get Started with QnA Maker and Azure Bot Service



